
Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.3 or any later version published by the Free Software Foundation; with the Invariant Sections being “Funding Free Software”, the Front-Cover Texts being (a) (see below), and with the Back-Cover Texts being (b) (see below). A copy of the license is included in the section entitled “GNU Free Documentation License”.

(a) The FSF’s Front-Cover Text is:
A GNU Manual

(b) The FSF’s Back-Cover Text is:
You have freedom to copy and modify this GNU Manual, like GNU software. Copies published by the Free Software Foundation raise funds for GNU development.
# Short Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1 Contributing to GCC Development</td>
<td>3</td>
</tr>
<tr>
<td>2 GCC and Portability</td>
<td>5</td>
</tr>
<tr>
<td>3 Interfacing to GCC Output</td>
<td>7</td>
</tr>
<tr>
<td>4 The GCC low-level runtime library</td>
<td>9</td>
</tr>
<tr>
<td>5 Language Front Ends in GCC</td>
<td>59</td>
</tr>
<tr>
<td>6 Source Tree Structure and Build System</td>
<td>61</td>
</tr>
<tr>
<td>7 Testsuites</td>
<td>77</td>
</tr>
<tr>
<td>8 Option specification files</td>
<td>103</td>
</tr>
<tr>
<td>9 Passes and Files of the Compiler</td>
<td>111</td>
</tr>
<tr>
<td>10 RTL Representation</td>
<td>125</td>
</tr>
<tr>
<td>11 GENERIC</td>
<td>177</td>
</tr>
<tr>
<td>12 GIMPLE</td>
<td>221</td>
</tr>
<tr>
<td>13 Analysis and Optimization of GIMPLE tuples</td>
<td>257</td>
</tr>
<tr>
<td>14 Analysis and Representation of Loops</td>
<td>269</td>
</tr>
<tr>
<td>15 Control Flow Graph</td>
<td>279</td>
</tr>
<tr>
<td>16 Machine Descriptions</td>
<td>289</td>
</tr>
<tr>
<td>17 Target Description Macros and Functions</td>
<td>393</td>
</tr>
<tr>
<td>18 Host Configuration</td>
<td>559</td>
</tr>
<tr>
<td>19 Makefile Fragments</td>
<td>563</td>
</tr>
<tr>
<td>20 collect2</td>
<td>567</td>
</tr>
<tr>
<td>21 Standard Header File Directories</td>
<td>569</td>
</tr>
<tr>
<td>22 Memory Management and Type Information</td>
<td>571</td>
</tr>
<tr>
<td>23 Plugins</td>
<td>579</td>
</tr>
<tr>
<td>24 Link Time Optimization</td>
<td>587</td>
</tr>
<tr>
<td>Funding Free Software</td>
<td>595</td>
</tr>
<tr>
<td>The GNU Project and GNU/Linux</td>
<td>597</td>
</tr>
<tr>
<td>GNU General Public License</td>
<td>599</td>
</tr>
<tr>
<td>GNU Free Documentation License</td>
<td>611</td>
</tr>
<tr>
<td>Contributors to GCC</td>
<td>619</td>
</tr>
<tr>
<td>Option Index</td>
<td>635</td>
</tr>
<tr>
<td>Concept Index</td>
<td>637</td>
</tr>
</tbody>
</table>
Table of Contents

Introduction ................................................. 1

1 Contributing to GCC Development ............... 3

2 GCC and Portability ................................. 5

3 Interfacing to GCC Output .......................... 7

4 The GCC low-level runtime library .......... 9
  4.1 Routines for integer arithmetic ............... 9
  4.1.1 Arithmetic functions ........................ 9
  4.1.2 Comparison functions ....................... 10
  4.1.3 Trapping arithmetic functions ............. 11
  4.1.4 Bit operations .............................. 11
  4.2 Routines for floating point emulation .. 12
  4.2.1 Arithmetic functions ....................... 12
  4.2.2 Conversion functions ....................... 13
  4.2.3 Comparison functions ....................... 15
  4.2.4 Other floating-point functions .......... 16
  4.3 Routines for decimal floating point emulation 16
  4.3.1 Arithmetic functions ....................... 17
  4.3.2 Conversion functions ....................... 17
  4.3.3 Comparison functions ....................... 20
  4.4 Routines for fixed-point fractional emulation 22
  4.4.1 Arithmetic functions ....................... 22
  4.4.2 Comparison functions ....................... 30
  4.4.3 Conversion functions ....................... 30
  4.5 Language-independent routines for exception handling 57
  4.6 Miscellaneous runtime library routines .... 57
  4.6.1 Cache control functions ................... 57
  4.6.2 Split stack functions and variables ...... 57

5 Language Front Ends in GCC ................. 59

6 Source Tree Structure and Build System .... 61
  6.1 Configure Terms and History .................. 61
  6.2 Top Level Source Directory .................... 61
  6.3 The ‘gcc’ Subdirectory ......................... 63
    6.3.1 Subdirectories of ‘gcc’ ................. 63
    6.3.2 Configuration in the ‘gcc’ Directory .. 64
      6.3.2.1 Scripts Used by ‘configure’ ...... 64
6.3.2.2 The 'config.build'; 'config.host'; and 'config.gcc'
Files .................................................. 64
6.3.2.3 Files Created by configure ...................... 65
6.3.3 Build System in the 'gcc' Directory ............... 65
6.3.4 Makefile Targets .................................. 65
6.3.5 Library Source Files and Headers under the 'gcc'
Directory .................................................. 68
6.3.6 Headers Installed by GCC .......................... 68
6.3.7 Build System in the 'gcc' Directory ............... 69
6.3.8 Anatomy of a Language Front End ................. 71
6.3.8.1 The Front End 'language' Directory ............ 72
6.3.8.2 The Front End 'config-lang.in' File ............ 73
6.3.8.3 The Front End 'Make-lang.in' File .............. 74
6.3.9 Anatomy of a Target Back End ..................... 75

7 Testsuites ............................................. 77
7.1 Idioms Used in Testsuite Code ....................... 77
7.2 Directives used within DejaGnu tests ............... 78
7.2.1 Syntax and Descriptions of test directives ....... 78
7.2.1.1 Specify how to build the test ................. 78
7.2.1.2 Specify additional compiler options .......... 79
7.2.1.3 Modify the test timeout value ................. 79
7.2.1.4 Skip a test for some targets ................. 79
7.2.1.5 Expect a test to fail for some targets ...... 80
7.2.1.6 Expect the test executable to fail .......... 80
7.2.1.7 Verify compiler messages .................... 80
7.2.1.8 Verify output of the test executable ......... 81
7.2.1.9 Specify additional files for a test .......... 81
7.2.1.10 Add checks at the end of a test ............. 81
7.2.2 Selecting targets to which a test applies ....... 81
7.2.3 Keywords describing target attributes .......... 82
7.2.3.1 Data type sizes ................................ 82
7.2.3.2 Fortran-specific attributes .................... 83
7.2.3.3 Vector-specific attributes .................... 83
7.2.3.4 Thread Local Storage attributes ............... 85
7.2.3.5 Decimal floating point attributes ............. 85
7.2.3.6 ARM-specific attributes ....................... 85
7.2.3.7 MIPS-specific attributes ....................... 86
7.2.3.8 PowerPC-specific attributes ................... 86
7.2.3.9 Other hardware attributes ..................... 87
7.2.3.10 Environment attributes ....................... 88
7.2.3.11 Other attributes ................................ 89
7.2.3.12 Local to tests in gcc.target/i386 .......... 91
7.2.3.13 Local to tests in gcc.target/spu/ea .......... 91
7.2.3.14 Local to tests in gcc.test-framework ....... 91
### 7.2.6 Commands for use in dg-final

- 7.2.6.1 Scan a particular file
- 7.2.6.2 Scan the assembly output
- 7.2.6.3 Scan optimization dump files
- 7.2.6.4 Verify that an output files exists or not
- 7.2.6.5 Check for LTO tests
- 7.2.6.6 Checks for gcov tests
- 7.2.6.7 Clean up generated test files

### 7.3 Ada Language Testsuites

### 7.4 C Language Testsuites

### 7.5 The Java library testsuites.

### 7.6 Support for testing link-time optimizations

### 7.7 Support for testing gcov

### 7.8 Support for testing profile-directed optimizations

### 7.9 Support for testing binary compatibility

### 7.10 Support for torture testing using multiple options

---

### 8 Option specification files

- 8.1 Option file format
- 8.2 Option properties

### 9 Passes and Files of the Compiler

- 9.1 Parsing pass
- 9.2 Gimplification pass
- 9.3 Pass manager
- 9.4 Tree SSA passes
- 9.5 RTL passes

### 10 RTL Representation

- 10.1 RTL Object Types
- 10.2 RTL Classes and Formats
- 10.3 Access to Operands
- 10.4 Access to Special Operands
- 10.5 Flags in an RTL Expression
- 10.6 Machine Modes
- 10.7 Constant Expression Types
- 10.8 Registers and Memory
- 10.9 RTL Expressions for Arithmetic
- 10.10 Comparison Operations
- 10.11 Bit-Fields
- 10.12 Vector Operations
- 10.13 Conversions
- 10.14 Declarations
- 10.15 Side Effect Expressions
- 10.16 Embedded Side-Effects on Addresses
11 GENERIC ............................................ 177
  11.1 Deficiencies ........................................... 177
  11.2 Overview ............................................. 177
    11.2.1 Trees ........................................... 178
    11.2.2 Identifiers ....................................... 179
    11.2.3 Containers ......................................... 179
  11.3 Types .................................................. 179
  11.4 Declarations .......................................... 184
    11.4.1 Working with declarations ......................... 184
    11.4.2 Internal structure ................................ 186
      11.4.2.1 Current structure hierarchy ................... 186
      11.4.2.2 Adding new DECL node types .................... 187
  11.5 Attributes in trees .................................... 188
  11.6 Expressions ........................................... 189
    11.6.1 Constant expressions ............................ 189
    11.6.2 References to storage ............................. 191
    11.6.3 Unary and Binary Expressions .................... 192
    11.6.4 Vectors ........................................... 199
  11.7 Statements ........................................... 200
    11.7.1 Basic Statements .................................. 201
    11.7.2 Blocks ............................................ 202
    11.7.3 Statement Sequences ................................ 203
    11.7.4 Empty Statements ................................... 203
    11.7.5 Jumps ............................................... 203
    11.7.6 Cleanups ........................................... 203
    11.7.7 OpenMP ............................................ 204
  11.8 Functions ............................................. 206
    11.8.1 Function Basics .................................... 206
    11.8.2 Function Properties ................................ 207
  11.9 Language-dependent trees ............................... 208
  11.10 C and C++ Trees ....................................... 209
    11.10.1 Types for C++ ..................................... 209
    11.10.2 Namespaces ......................................... 211
    11.10.3 Classes ............................................ 212
    11.10.4 Functions for C++ .................................. 214
    11.10.5 Statements for C++ ................................ 217
      11.10.5.1 Statements .................................... 217
    11.10.6 C++ Expressions .................................... 220
  11.11 Java Trees ............................................ 220
12 GIMPLE ........................................ 221

12.1 Tuple representation .................................. 222
  12.1.1 gimple_statement_base (gsbase) .................. 222
  12.1.2 gimple_statement_with_ops ...................... 223
  12.1.3 gimple_statement_with_memory_ops .............. 223

12.2 GIMPLE instruction set ................................ 225

12.3 Exception Handling .................................. 225

12.4 Temporaries .......................................... 226

12.5 Operands ............................................. 226
  12.5.1 Compound Expressions ............................ 227
  12.5.2 Compound Lvalues ............................... 227
  12.5.3 Conditional Expressions ......................... 227
  12.5.4 Logical Operators ............................... 227
  12.5.5 Manipulating operands ............................ 228
  12.5.6 Operand vector allocation ....................... 228
  12.5.7 Operand validation ................................ 229
  12.5.8 Statement validation ............................. 230

12.6 Manipulating GIMPLE statements ....................... 230
  12.6.1 Common accessors .................................. 230

12.7 Tuple specific accessors ............................. 233
  12.7.1 GIMPLE_ASM ....................................... 233
  12.7.2 GIMPLE_ASSIGN .................................... 234
  12.7.3 GIMPLE_BIND ...................................... 235
  12.7.4 GIMPLE_CALL ...................................... 236
  12.7.5 GIMPLE_CATCH .................................... 237
  12.7.6 GIMPLE_COND ..................................... 238
  12.7.7 GIMPLE_DEBUG .................................... 239
  12.7.8 GIMPLE_EH_FILTER ................................ 240
  12.7.9 GIMPLE_LABEL ..................................... 241
  12.7.10 GIMPLE_NOP ....................................... 241
  12.7.11 GIMPLE_OMP_ATOMIC_LOAD ....................... 241
  12.7.12 GIMPLE_OMP_ATOMIC_STORE ...................... 241
  12.7.13 GIMPLE_OMP_CONTINUE ............................ 242
  12.7.14 GIMPLE_OMP_CRITICAL ........................... 242
  12.7.15 GIMPLE_OMP_FOR ................................. 243
  12.7.16 GIMPLE_OMP_MASTER ............................. 244
  12.7.17 GIMPLE_OMP_ORDERED ............................. 244
  12.7.18 GIMPLE_OMP_PARALLEL ........................... 244
  12.7.19 GIMPLE_OMP_RETURN .............................. 245
  12.7.20 GIMPLE_OMP_SECTION ............................ 246
  12.7.21 GIMPLE_OMP_SECTIONS ............................ 246
  12.7.22 GIMPLE_OMP_SINGLE ............................. 246
  12.7.23 GIMPLE_PHI ....................................... 247
  12.7.24 GIMPLE_RESX ..................................... 247
  12.7.25 GIMPLE_RETURN .................................. 248
  12.7.26 GIMPLE_SWITCH ................................... 248
  12.7.27 GIMPLE_TRY ....................................... 249
  12.7.28 GIMPLE_WITH_CLEANUP_EXPR ..................... 249
12.8 GIMPLE sequences ........................................... 250
12.9 Sequence iterators ........................................... 251
12.10 Adding a new GIMPLE statement code ................. 254
12.11 Statement and operand traversals ....................... 254

13 Analysis and Optimization of GIMPLE tuples ............ 257
  13.1 Annotations .................................................. 257
  13.2 SSA Operands ............................................... 257
    13.2.1 Operand Iterators And Access Routines .......... 259
    13.2.2 Immediate Uses .................................... 261
  13.3 Static Single Assignment .................................. 263
    13.3.1 Preserving the SSA form ......................... 264
    13.3.2 Preserving the virtual SSA form ................ 265
    13.3.3 Examining SSA_NAME nodes ....................... 266
    13.3.4 Walking use-def chains .......................... 266
    13.3.5 Walking the dominator tree ...................... 266
  13.4 Alias analysis ............................................. 267
  13.5 Memory model .............................................. 268

14 Analysis and Representation of Loops .................... 269
  14.1 Loop representation ...................................... 269
  14.2 Loop querying ............................................ 271
  14.3 Loop manipulation ....................................... 272
  14.4 Loop-closed SSA form .................................... 272
  14.5 Scalar evolutions ....................................... 273
  14.6 IV analysis on RTL ....................................... 274
  14.7 Number of iterations analysis .......................... 274
  14.8 Data Dependency Analysis ............................... 275
  14.9 Linear loop transformations framework ............... 277
  14.10 Omega a solver for linear programming problems .... 277

15 Control Flow Graph ........................................... 279
  15.1 Basic Blocks .............................................. 279
  15.2 Edges ..................................................... 280
  15.3 Profile information ...................................... 283
  15.4 Maintaining the CFG .................................... 284
  15.5 Liveness information ................................... 286
# Machine Descriptions

16.1 Overview of How the Machine Description is Used ........................................ 289
16.2 Everything about Instruction Patterns ....................................................... 289
16.3 Example of `define_insn` .................................................................................. 290
16.4 RTL Template .................................................................................................... 291
16.5 Output Templates and Operand Substitution ................................................... 295
16.6 C Statements for Assembler Output ..................................................................... 296
16.7 Predicates .......................................................................................................... 297
  16.7.1 Machine-Independent Predicates ................................................................. 298
  16.7.2 Defining Machine-Specific Predicates ......................................................... 300
16.8 Operand Constraints ......................................................................................... 302
  16.8.1 Simple Constraints ....................................................................................... 302
  16.8.2 Multiple Alternative Constraints ................................................................. 306
  16.8.3 Register Class Preferences ......................................................................... 307
  16.8.4 Constraint Modifier Characters ................................................................... 307
  16.8.5 Constraints for Particular Machines ............................................................. 308
  16.8.6 Disable insn alternatives using the `enabled` attribute ............................... 328
  16.8.7 Defining Machine-Specific Constraints ....................................................... 329
  16.8.8 Testing constraints from C ........................................................................... 331
16.9 Standard Pattern Names For Generation ......................................................... 332
16.10 When the Order of Patterns Matters .............................................................. 357
16.11 Interdependence of Patterns ............................................................................ 357
16.12 Defining Jump Instruction Patterns .................................................................. 357
16.13 Defining Looping Instruction Patterns ........................................................... 358
16.14 Canonicalization of Instructions ...................................................................... 360
16.15 Defining RTL Sequences for Code Generation ................................................. 361
16.16 Defining How to Split Instructions ................................................................... 364
16.17 Including Patterns in Machine Descriptions ................................................ 367
  16.17.1 RTL Generation Tool Options for Directory Search ................................. 367
16.18 Machine-Specific Peephole Optimizers ........................................................... 368
  16.18.1 RTL to Text Peephole Optimizers ............................................................... 368
  16.18.2 RTL to RTL Peephole Optimizers ............................................................... 370
16.19 Instruction Attributes ..................................................................................... 371
  16.19.1 Defining Attributes and their Values ......................................................... 372
  16.19.2 Attribute Expressions ............................................................................... 373
  16.19.3 Assigning Attribute Values to Insns ......................................................... 376
  16.19.4 Example of Attribute Specifications ......................................................... 377
  16.19.5 Computing the Length of an Insn ............................................................... 378
  16.19.6 Constant Attributes ................................................................................. 379
  16.19.7 Delay Slot Scheduling .............................................................................. 379
  16.19.8 Specifying processor pipeline description ................................................. 380
16.20 Conditional Execution ..................................................................................... 386
16.21 Constant Definitions ....................................................................................... 387
16.22 Iterators ............................................................................................................ 389
  16.22.1 Mode Iterators ............................................................................................ 389
    16.22.1.1 Defining Mode Iterators ..................................................................... 389
    16.22.1.2 Substitution in Mode Iterators ............................................................. 390
    16.22.1.3 Mode Iterator Examples .................................................................... 390
17 Target Description Macros and Functions

17.1 The Global targetm Variable ............................................ 393
17.2 Controlling the Compilation Driver, 'gcc' .......................... 393
17.3 Run-time Target Specification ......................................... 400
17.4 Defining data structures for per-function information ............. 403
17.5 Storage Layout ............................................................ 404
17.6 Layout of Source Language Data Types ............................. 413
17.7 Register Usage ............................................................ 418
  17.7.1 Basic Characteristics of Registers ............................... 418
  17.7.2 Order of Allocation of Registers ............................... 420
  17.7.3 How Values Fit in Registers .................................... 421
  17.7.4 Handling Leaf Functions ........................................ 423
  17.7.5 Registers That Form a Stack .................................... 424
17.8 Register Classes .......................................................... 425
17.9 Obsolete Macros for Defining Constraints .......................... 433
17.10 Stack Layout and Calling Conventions ............................ 435
  17.10.1 Basic Stack Layout .............................................. 436
  17.10.2 Exception Handling Support .................................... 439
  17.10.3 Specifying How Stack Checking is Done ....................... 441
  17.10.4 Registers That Address the Stack Frame ...................... 443
  17.10.5 Eliminating Frame Pointer and Arg Pointer .................. 445
  17.10.6 Passing Function Arguments on the Stack ...................... 447
  17.10.7 Passing Arguments in Registers ................................ 449
  17.10.8 How Scalar Function Values Are Returned .................... 455
  17.10.9 How Large Values Are Returned ................................ 457
  17.10.10 Caller-Saves Register Allocation .............................. 458
  17.10.11 Function Entry and Exit ....................................... 458
  17.10.12 Generating Code for Profiling ................................ 462
  17.10.13 Permitting tail calls ......................................... 463
  17.10.14 Stack smashing protection .................................... 463
17.11 Implementing the Varargs Macros .................................. 464
17.12 Trampolines for Nested Functions .................................. 466
17.13 Implicit Calls to Library Routines ................................ 468
17.14 Addressing Modes ...................................................... 469
17.15 Anchored Addresses .................................................... 475
17.16 Condition Code Status .................................................. 476
  17.16.1 Representation of condition codes using (cc0) ............... 477
  17.16.2 Representation of condition codes using registers .......... 478
  17.16.3 Macros to control conditional execution ...................... 480
17.17 Describing Relative Costs of Operations .......................... 480
17.18 Adjusting the Instruction Scheduler ................................ 485
17.19 Dividing the Output into Sections (Texts, Data, ...) ............. 491
17.20 Position Independent Code .......................................... 496
17.21 Defining the Output Assembler Language .......................... 497
  17.21.1 The Overall Framework of an Assembler File ................. 497
17.21.2 Output of Data ............................................. 500
17.21.3 Output of Uninitialized Variables ................. 503
17.21.4 Output and Generation of Labels ................. 505
17.21.5 How Initialization Functions Are Handled .......... 512
17.21.6 Macros Controlling Initialization Routines ........ 514
17.21.7 Output of Assembler Instructions .................. 516
17.21.8 Output of Dispatch Tables .......................... 519
17.21.9 Assembler Commands for Exception Regions ......... 521
17.21.10 Assembler Commands for Alignment ............... 523
17.22 Controlling Debugging Information Format .......... 525
17.22.1 Macros Affecting All Debugging Formats .......... 525
17.22.2 Specific Options for DBX Output .................... 526
17.22.3 Open-Ended Hooks for DBX Format ................. 528
17.22.4 File Names in DBX Format .......................... 528
17.22.5 Macros for SDB and DWARF Output ................. 529
17.22.6 Macros for VMS Debug Format ...................... 531
17.23 Cross Compilation and Floating Point ............... 531
17.24 Mode Switching Instructions .......................... 533
17.25 Defining target-specific uses of __attribute__ ........ 534
17.26 Emulating TLS ........................................ 537
17.27 Defining coprocessor specifics for MIPS targets ..... 538
17.28 Parameters for Precompiled Header Validity Checking .... 538
17.29 C++ ABI parameters ................................. 539
17.30 Adding support for named address spaces ............ 540
17.31 Miscellaneous Parameters ............................ 542

18 Host Configuration ........................................ 559
18.1 Host Common ........................................ 559
18.2 Host Filesystem ....................................... 560
18.3 Host Misc .............................................. 561

19 Makefile Fragments ..................................... 563
19.1 Target Makefile Fragments ............................ 563
19.2 Host Makefile Fragments ............................... 565

20 collect2 ............................................. 567

21 Standard Header File Directories ................. 569

22 Memory Management and Type Information ............. 571
22.1 The Inside of a GTY(0) .............................. 571
22.2 Marking Roots for the Garbage Collector .............. 576
22.3 Source Files Containing Type Information .......... 576
22.4 How to invoke the garbage collector ................. 577
22.5 Troubleshooting the garbage collector .............. 577
Introduction

This manual documents the internals of the GNU compilers, including how to port them to new targets and some information about how to write front ends for new languages. It corresponds to the compilers (GCC) version 4.6.2. The use of the GNU compilers is documented in a separate manual. See Section “Introduction” in Using the GNU Compiler Collection (GCC).

This manual is mainly a reference manual rather than a tutorial. It discusses how to contribute to GCC (see Chapter 1 [Contributing], page 3), the characteristics of the machines supported by GCC as hosts and targets (see Chapter 2 [Portability], page 5), how GCC relates to the ABIs on such systems (see Chapter 3 [Interface], page 7), and the characteristics of the languages for which GCC front ends are written (see Chapter 5 [Languages], page 59). It then describes the GCC source tree structure and build system, some of the interfaces to GCC front ends, and how support for a target system is implemented in GCC.

Additional tutorial information is linked to from http://gcc.gnu.org/readings.html.
1 Contributing to GCC Development

If you would like to help pretest GCC releases to assure they work well, current development sources are available by SVN (see http://gcc.gnu.org/svn.html). Source and binary snapshots are also available for FTP; see http://gcc.gnu.org/snapshots.html.

If you would like to work on improvements to GCC, please read the advice at these URLs:

http://gcc.gnu.org/contribute.html
http://gcc.gnu.org/contributewhy.html

for information on how to make useful contributions and avoid duplication of effort. Suggested projects are listed at http://gcc.gnu.org/projects/.
2 GCC and Portability

GCC itself aims to be portable to any machine where int is at least a 32-bit type. It aims to target machines with a flat (non-segmented) byte addressed data address space (the code address space can be separate). Target ABIs may have 8, 16, 32 or 64-bit int type. char can be wider than 8 bits.

GCC gets most of the information about the target machine from a machine description which gives an algebraic formula for each of the machine’s instructions. This is a very clean way to describe the target. But when the compiler needs information that is difficult to express in this fashion, ad-hoc parameters have been defined for machine descriptions. The purpose of portability is to reduce the total work needed on the compiler; it was not of interest for its own sake.

GCC does not contain machine dependent code, but it does contain code that depends on machine parameters such as endianness (whether the most significant byte has the highest or lowest address of the bytes in a word) and the availability of autoincrement addressing. In the RTL-generation pass, it is often necessary to have multiple strategies for generating code for a particular kind of syntax tree, strategies that are usable for different combinations of parameters. Often, not all possible cases have been addressed, but only the common ones or only the ones that have been encountered. As a result, a new target may require additional strategies. You will know if this happens because the compiler will call abort. Fortunately, the new strategies can be added in a machine-independent fashion, and will affect only the target machines that need them.
3 Interfacing to GCC Output

GCC is normally configured to use the same function calling convention normally in use on the target system. This is done with the machine-description macros described (see Chapter 17 [Target Macros], page 393).

However, returning of structure and union values is done differently on some target machines. As a result, functions compiled with PCC returning such types cannot be called from code compiled with GCC, and vice versa. This does not cause trouble often because few Unix library routines return structures or unions.

GCC code returns structures and unions that are 1, 2, 4 or 8 bytes long in the same registers used for \texttt{int} or \texttt{double} return values. (GCC typically allocates variables of such types in registers also.) Structures and unions of other sizes are returned by storing them into an address passed by the caller (usually in a register). The target hook \texttt{TARGET_STRUCT_VALUE_RTX} tells GCC where to pass this address.

By contrast, PCC on most target machines returns structures and unions of any size by copying the data into an area of static storage, and then returning the address of that storage as if it were a pointer value. The caller must copy the data from that memory area to the place where the value is wanted. This is slower than the method used by GCC, and fails to be reentrant.

On some target machines, such as RISC machines and the 80386, the standard system convention is to pass to the subroutine the address of where to return the value. On these machines, GCC has been configured to be compatible with the standard compiler, when this method is used. It may not be compatible for structures of 1, 2, 4 or 8 bytes.

GCC uses the system’s standard convention for passing arguments. On some machines, the first few arguments are passed in registers; in others, all are passed on the stack. It would be possible to use registers for argument passing on any machine, and this would probably result in a significant speedup. But the result would be complete incompatibility with code that follows the standard convention. So this change is practical only if you are switching to GCC as the sole C compiler for the system. We may implement register argument passing on certain machines once we have a complete GNU system so that we can compile the libraries with GCC.

On some machines (particularly the SPARC), certain types of arguments are passed “by invisible reference”. This means that the value is stored in memory, and the address of the memory location is passed to the subroutine.

If you use \texttt{longjmp}, beware of automatic variables. ISO C says that automatic variables that are not declared \texttt{volatile} have undefined values after a \texttt{longjmp}. And this is all GCC promises to do, because it is very difficult to restore register variables correctly, and one of GCC’s features is that it can put variables in registers without your asking it to.
4 The GCC low-level runtime library

GCC provides a low-level runtime library, ‘libgcc.a’ or ‘libgcc_s.so.1’ on some platforms. GCC generates calls to routines in this library automatically, whenever it needs to perform some operation that is too complicated to emit inline code for.

Most of the routines in libgcc handle arithmetic operations that the target processor cannot perform directly. This includes integer multiply and divide on some machines, and all floating-point and fixed-point operations on other machines. libgcc also includes routines for exception handling, and a handful of miscellaneous operations.

Some of these routines can be defined in mostly machine-independent C. Others must be hand-written in assembly language for each processor that needs them.

GCC will also generate calls to C library routines, such as memcpy and memset, in some cases. The set of routines that GCC may possibly use is documented in Section “Other Builtins” in Using the GNU Compiler Collection (GCC).

These routines take arguments and return values of a specific machine mode, not a specific C type. See Section 10.6 [Machine Modes], page 137, for an explanation of this concept. For illustrative purposes, in this chapter the floating point type float is assumed to correspond to SFmode; double to DFmode; and long double to both TFmode and XFmode. Similarly, the integer types int and unsigned int correspond to SImode; long and unsigned long to DImode; and long long and unsigned long long to TImode.

4.1 Routines for integer arithmetic

The integer arithmetic routines are used on platforms that don’t provide hardware support for arithmetic operations on some modes.

4.1.1 Arithmetic functions

int __ashlsi3 (int a, int b)  
long __ashldi3 (long a, int b)  
long long __ashlti3 (long long a, int b)  

These functions return the result of shifting a left by b bits.

int __ashrsi3 (int a, int b)  
long __ashrdi3 (long a, int b)  
long long __ashrti3 (long long a, int b)  

These functions return the result of arithmetically shifting a right by b bits.

int __divsi3 (int a, int b)  
long __divdi3 (long a, long b)  
long long __divti3 (long long a, long long b)  

These functions return the quotient of the signed division of a and b.

int __lshrsi3 (int a, int b)  
long __lshrdi3 (long a, int b)  
long long __lshrti3 (long long a, int b)  

These functions return the result of logically shifting a right by b bits.
int __modsi3 (int a, int b)  
long __moddi3 (long a, long b)  
long long __modti3 (long long a, long long b)

These functions return the remainder of the signed division of a and b.

int __mulsi3 (int a, int b)  
long __muldi3 (long a, long b)  
long long __multi3 (long long a, long long b)

These functions return the product of a and b.

long __negdi2 (long a)  
long long __negti2 (long long a)

These functions return the negation of a.

unsigned int __udivsi3 (unsigned int a, unsigned int b)  
unsigned long __udivdi3 (unsigned long a, unsigned long b)  
unsigned long long __udivti3 (unsigned long long a, unsigned long long b)

These functions return the quotient of the unsigned division of a and b.

unsigned long __udivmoddi3 (unsigned long a, unsigned long b, unsigned long *c)  
unsigned long long __udivti3 (unsigned long long a, unsigned long long b, unsigned long long *c)

These functions calculate both the quotient and remainder of the unsigned division of a and b. The return value is the quotient, and the remainder is placed in variable pointed to by c.

unsigned int __umodsi3 (unsigned int a, unsigned int b)  
unsigned long __umoddi3 (unsigned long a, unsigned long b)  
unsigned long long __umodti3 (unsigned long long a, unsigned long long b)

These functions return the remainder of the unsigned division of a and b.

4.1.2 Comparison functions

The following functions implement integral comparisons. These functions implement a low-level compare, upon which the higher level comparison operators (such as less than and greater than or equal to) can be constructed. The returned values lie in the range zero to two, to allow the high-level operators to be implemented by testing the returned result using either signed or unsigned comparison.

int __cmpdi2 (long a, long b)  
int __cmpti2 (long long a, long long b)

These functions perform a signed comparison of a and b. If a is less than b, they return 0; if a is greater than b, they return 2; and if a and b are equal they return 1.

int __ucmpdi2 (unsigned long a, unsigned long b)  
int __ucmpti2 (unsigned long long a, unsigned long long b)

These functions perform an unsigned comparison of a and b. If a is less than b, they return 0; if a is greater than b, they return 2; and if a and b are equal they return 1.
4.1.3 Trapping arithmetic functions

The following functions implement trapping arithmetic. These functions call the libc function abort upon signed arithmetic overflow.

\[
\begin{align*}
\text{int } & \text{ } _{-}\text{absvsi2 (int a)} \quad \text{[Runtime Function]} \\
\text{long } & \text{ } _{-}\text{absvdi2 (long a)} \quad \text{[Runtime Function]}
\end{align*}
\]

These functions return the absolute value of \(a\).

\[
\begin{align*}
\text{int } & \text{ } _{-}\text{addvsi3 (int a, int b)} \quad \text{[Runtime Function]} \\
\text{long } & \text{ } _{-}\text{addvdi3 (long a, long b)} \quad \text{[Runtime Function]}
\end{align*}
\]

These functions return the sum of \(a\) and \(b\); that is \(a + b\).

\[
\begin{align*}
\text{int } & \text{ } _{-}\text{mulvsi3 (int a, int b)} \quad \text{[Runtime Function]} \\
\text{long } & \text{ } _{-}\text{mulvdi3 (long a, long b)} \quad \text{[Runtime Function]}
\end{align*}
\]

The functions return the product of \(a\) and \(b\); that is \(a \times b\).

\[
\begin{align*}
\text{int } & \text{ } _{-}\text{negvsi2 (int a)} \quad \text{[Runtime Function]} \\
\text{long } & \text{ } _{-}\text{negvdi2 (long a)} \quad \text{[Runtime Function]}
\end{align*}
\]

These functions return the negation of \(a\); that is \(-a\).

\[
\begin{align*}
\text{int } & \text{ } _{-}\text{subvsi3 (int a, int b)} \quad \text{[Runtime Function]} \\
\text{long } & \text{ } _{-}\text{subvdi3 (long a, long b)} \quad \text{[Runtime Function]}
\end{align*}
\]

These functions return the difference between \(b\) and \(a\); that is \(a - b\).

4.1.4 Bit operations

\[
\begin{align*}
\text{int } & \text{ } _{-}\text{clzsi2 (int a)} \quad \text{[Runtime Function]} \\
\text{int } & \text{ } _{-}\text{clzdi2 (long a)} \quad \text{[Runtime Function]} \\
\text{int } & \text{ } _{-}\text{clzti2 (long long a)} \quad \text{[Runtime Function]}
\end{align*}
\]

These functions return the number of leading 0-bits in \(a\), starting at the most significant bit position. If \(a\) is zero, the result is undefined.

\[
\begin{align*}
\text{int } & \text{ } _{-}\text{ctzsi2 (int a)} \quad \text{[Runtime Function]} \\
\text{int } & \text{ } _{-}\text{ctzdi2 (long a)} \quad \text{[Runtime Function]} \\
\text{int } & \text{ } _{-}\text{ctzti2 (long long a)} \quad \text{[Runtime Function]}
\end{align*}
\]

These functions return the number of trailing 0-bits in \(a\), starting at the least significant bit position. If \(a\) is zero, the result is undefined.

\[
\begin{align*}
\text{int } & \text{ } _{-}\text{ffsdi2 (long a)} \quad \text{[Runtime Function]} \\
\text{int } & \text{ } _{-}\text{ffsti2 (long long a)} \quad \text{[Runtime Function]}
\end{align*}
\]

These functions return the index of the least significant 1-bit in \(a\), or the value zero if \(a\) is zero. The least significant bit is index one.

\[
\begin{align*}
\text{int } & \text{ } _{-}\text{paritysi2 (int a)} \quad \text{[Runtime Function]} \\
\text{int } & \text{ } _{-}\text{paritydi2 (long a)} \quad \text{[Runtime Function]} \\
\text{int } & \text{ } _{-}\text{parityti2 (long long a)} \quad \text{[Runtime Function]}
\end{align*}
\]

These functions return the value zero if the number of bits set in \(a\) is even, and the value one otherwise.
int __popcountsi2 (int a)  
int __popcountdi2 (long a)  
int __popcountti2 (long long a)  

These functions return the number of bits set in a.

int32_t __bswapsi2 (int32_t a)  
int64_t __bswapdi2 (int64_t a)  

These functions return the a byteswapped.

4.2 Routines for floating point emulation

The software floating point library is used on machines which do not have hardware support
for floating point. It is also used whenever ‘-msoft-float’ is used to disable generation of
floating point instructions. (Not all targets support this switch.)

For compatibility with other compilers, the floating point emulation routines can be
renamed with the DECLARE_LIBRARY_RENAMES macro (see Section 17.13 [Library Calls],
page 468). In this section, the default names are used.

Presently the library does not support XFmode, which is used for long double on some
architectures.

4.2.1 Arithmetic functions

float __addsf3 (float a, float b)  
double __adddf3 (double a, double b)  
long double __addtf3 (long double a, long double b)  
long double __addxf3 (long double a, long double b)  

These functions return the sum of a and b.

float __subsf3 (float a, float b)  
double __subdf3 (double a, double b)  
long double __subtf3 (long double a, long double b)  
long double __subxf3 (long double a, long double b)  

These functions return the difference between b and a; that is, a – b.

float __mulsf3 (float a, float b)  
double __muldf3 (double a, double b)  
long double __multf3 (long double a, long double b)  
long double __mulxf3 (long double a, long double b)  

These functions return the product of a and b.

float __divsf3 (float a, float b)  
double __divdf3 (double a, double b)  
long double __divtf3 (long double a, long double b)  
long double __divxf3 (long double a, long double b)  

These functions return the quotient of a and b; that is, a/b.

float __negsf2 (float a)  
double __negdf2 (double a)  
long double __negtf2 (long double a)
long double __negxf2 (long double a) [Runtime Function]
These functions return the negation of a. They simply flip the sign bit, so they can produce negative zero and negative NaN.

4.2.2 Conversion functions

double __extendsfdf2 (float a) [Runtime Function]
long double __extendsstf2 (float a) [Runtime Function]
long double __extendsxf2 (float a) [Runtime Function]
long double __extendstf2 (double a) [Runtime Function]
long double __extendxf2 (double a) [Runtime Function]
These functions extend a to the wider mode of their return type.

double __truncxfdf2 (long double a) [Runtime Function]
double __trunctfdf2 (long double a) [Runtime Function]
float __truncxfsf2 (long double a) [Runtime Function]
float __trunctfsf2 (long double a) [Runtime Function]
float __truncdfsf2 (double a) [Runtime Function]
These functions truncate a to the narrower mode of their return type, rounding toward zero.

int __fixsfsi (float a) [Runtime Function]
int __fixdfsi (double a) [Runtime Function]
int __fixtfsi (long double a) [Runtime Function]
int __fixxfsi (long double a) [Runtime Function]
These functions convert a to a signed integer, rounding toward zero.

long __fixsfdi (float a) [Runtime Function]
long __fixdfdi (double a) [Runtime Function]
long __fixtfdi (long double a) [Runtime Function]
long __fixxfdi (long double a) [Runtime Function]
These functions convert a to a signed long, rounding toward zero.

long long __fixsfti (float a) [Runtime Function]
long long __fixdfhti (double a) [Runtime Function]
long long __fixtfti (long double a) [Runtime Function]
long long __fixxffti (long double a) [Runtime Function]
These functions convert a to a signed long long, rounding toward zero.

unsigned int __fixunssfsi (float a) [Runtime Function]
unsigned int __fixunssdfsi (double a) [Runtime Function]
unsigned int __fixunstfsi (long double a) [Runtime Function]
unsigned int __fixunxxfsi (long double a) [Runtime Function]
These functions convert a to an unsigned integer, rounding toward zero. Negative values all become zero.

unsigned long __fixunssfdi (float a) [Runtime Function]
unsigned long __fixunsdfdi (double a) [Runtime Function]
unsigned long __fixunstfdi (long double a) [Runtime Function]
unsigned long __fixunsxfdi (long double a) [Runtime Function]

These functions convert a to an unsigned long, rounding toward zero. Negative values all become zero.

unsigned long long __fixunssfti (float a) [Runtime Function]
unsigned long long __fixunsdfti (double a) [Runtime Function]
unsigned long long __fixunstfti (long double a) [Runtime Function]
unsigned long long __fixunsxfti (long double a) [Runtime Function]

These functions convert a to an unsigned long long, rounding toward zero. Negative values all become zero.

float __floatsisf (int i) [Runtime Function]
double __floatsidf (int i) [Runtime Function]
long double __floatsitf (int i) [Runtime Function]
long double __floatsif (int i) [Runtime Function]

These functions convert i, a signed integer, to floating point.

float __floatdisf (long i) [Runtime Function]
double __floatdidxf (long i) [Runtime Function]
long double __floatdfxf (long i) [Runtime Function]
long double __floatdf (long i) [Runtime Function]

These functions convert i, a signed long, to floating point.

float __floattisf (long long i) [Runtime Function]
double __floattitf (long long i) [Runtime Function]
long double __floattitf (long long i) [Runtime Function]
long double __floattixf (long long i) [Runtime Function]

These functions convert i, a signed long long, to floating point.

float __floatunsisf (unsigned int i) [Runtime Function]
double __floatunsidf (unsigned int i) [Runtime Function]
long double __floatunsitf (unsigned int i) [Runtime Function]
long double __floatunsixf (unsigned int i) [Runtime Function]

These functions convert i, an unsigned integer, to floating point.

float __floatundisf (unsigned long i) [Runtime Function]
double __floatundidf (unsigned long i) [Runtime Function]
long double __floatunditf (unsigned long i) [Runtime Function]
long double __floatundixf (unsigned long i) [Runtime Function]

These functions convert i, an unsigned long, to floating point.

float __floatuntisf (unsigned long long i) [Runtime Function]
double __floatuntidf (unsigned long long i) [Runtime Function]
long double __floatuntitf (unsigned long long i) [Runtime Function]
long double __floatuntixf (unsigned long long i) [Runtime Function]

These functions convert i, an unsigned long long, to floating point.
4.2.3 Comparison functions

There are two sets of basic comparison functions.

```c
int __cmpsf2 (float a, float b)  [Runtime Function]
int __cmpdf2 (double a, double b) [Runtime Function]
int __cmptf2 (long double a, long double b) [Runtime Function]
```

These functions calculate \( a \leq b \). That is, if \( a \) is less than \( b \), they return \(-1\); if \( a \) is greater than \( b \), they return \(1\); and if \( a \) and \( b \) are equal they return \(0\). If either argument is NaN they return \(1\), but you should not rely on this; if NaN is a possibility, use one of the higher-level comparison functions.

```c
int __unordsf2 (float a, float b)  [Runtime Function]
int __unorddf2 (double a, double b) [Runtime Function]
int __unordtf2 (long double a, long double b) [Runtime Function]
```

These functions return a nonzero value if either argument is NaN, otherwise \(0\).

There is also a complete group of higher level functions which correspond directly to comparison operators. They implement the ISO C semantics for floating-point comparisons, taking NaN into account. Pay careful attention to the return values defined for each set.

Under the hood, all of these routines are implemented as

```
if (__unordXf2 (a, b))
  return E;
return __cmpXf2 (a, b);
```

where \( E \) is a constant chosen to give the proper behavior for NaN. Thus, the meaning of the return value is different for each set. Do not rely on this implementation; only the semantics documented below are guaranteed.

```c
int __eqsf2 (float a, float b)  [Runtime Function]
int __eqdf2 (double a, double b) [Runtime Function]
int __eqtf2 (long double a, long double b) [Runtime Function]
```

These functions return zero if neither argument is NaN, and \( a \) and \( b \) are equal.

```c
int __nesf2 (float a, float b)  [Runtime Function]
int __nedf2 (double a, double b) [Runtime Function]
int __netf2 (long double a, long double b) [Runtime Function]
```

These functions return a nonzero value if either argument is NaN, or if \( a \) and \( b \) are unequal.

```c
int __gesf2 (float a, float b)  [Runtime Function]
int __gedf2 (double a, double b) [Runtime Function]
int __getf2 (long double a, long double b) [Runtime Function]
```

These functions return a value greater than or equal to zero if neither argument is NaN, and \( a \) is greater than or equal to \( b \).

```c
int __ltsf2 (float a, float b)  [Runtime Function]
int __ltdf2 (double a, double b) [Runtime Function]
int __lttf2 (long double a, long double b) [Runtime Function]
```

These functions return a value less than zero if neither argument is NaN, and \( a \) is strictly less than \( b \).
int __lesf2 (float a, float b) [Runtime Function]
int __ledf2 (double a, double b) [Runtime Function]
int __letf2 (long double a, long double b) [Runtime Function]

These functions return a value less than or equal to zero if neither argument is NaN, and a is less than or equal to b.

int __gtsf2 (float a, float b) [Runtime Function]
int __gtdf2 (double a, double b) [Runtime Function]
int __gttf2 (long double a, long double b) [Runtime Function]

These functions return a value greater than zero if neither argument is NaN, and a is strictly greater than b.

### 4.2.4 Other floating-point functions

float __powisf2 (float a, int b) [Runtime Function]
double __powidf2 (double a, int b) [Runtime Function]
long double __powitf2 (long double a, int b) [Runtime Function]
long double __powixf2 (long double a, int b) [Runtime Function]

These functions convert raise a to the power b.

complex float __mulsc3 (float a, float b, float c, float d) [Runtime Function]
complex double __muldc3 (double a, double b, double c, double d) [Runtime Function]
complex long double __multc3 (long double a, long double b, long double c, long double d) [Runtime Function]
complex long double __mulxc3 (long double a, long double b, long double c, long double d) [Runtime Function]

These functions return the product of $a + ib$ and $c + id$, following the rules of C99 Annex G.

complex float __divsc3 (float a, float b, float c, float d) [Runtime Function]
complex double __divdc3 (double a, double b, double c, double d) [Runtime Function]
complex long double __divtc3 (long double a, long double b, long double c, long double d) [Runtime Function]
complex long double __divxc3 (long double a, long double b, long double c, long double d) [Runtime Function]

These functions return the quotient of $a + ib$ and $c + id$ (i.e., $(a + ib)/(c + id)$), following the rules of C99 Annex G.

### 4.3 Routines for decimal floating point emulation

The software decimal floating point library implements IEEE 754-2008 decimal floating point arithmetic and is only activated on selected targets.

The software decimal floating point library supports either DPD (Densely Packed Decimal) or BID (Binary Integer Decimal) encoding as selected at configure time.
4.3.1 Arithmetic functions

[Runtime Function] _Decimal32 __dpd_addsd3 (Decimal32 a, Decimal32 b)  
[Runtime Function] _Decimal32 __bid_addsd3 (Decimal32 a, Decimal32 b)  
[Runtime Function] _Decimal64 __dpd_adddd3 (Decimal64 a, Decimal64 b)  
[Runtime Function] _Decimal64 __bid_adddd3 (Decimal64 a, Decimal64 b)  
[Runtime Function] _Decimal128 __dpd_addtd3 (Decimal128 a, Decimal128 b)  
[Runtime Function] _Decimal128 __bid_addtd3 (Decimal128 a, Decimal128 b)  

These functions return the sum of a and b.

[Runtime Function] _Decimal32 __dpd_subsd3 (Decimal32 a, Decimal32 b)  
[Runtime Function] _Decimal32 __bid_subsd3 (Decimal32 a, Decimal32 b)  
[Runtime Function] _Decimal64 __dpd_subdd3 (Decimal64 a, Decimal64 b)  
[Runtime Function] _Decimal64 __bid_subdd3 (Decimal64 a, Decimal64 b)  
[Runtime Function] _Decimal128 __dpd_subtd3 (Decimal128 a, Decimal128 b)  
[Runtime Function] _Decimal128 __bid_subtd3 (Decimal128 a, Decimal128 b)  

These functions return the difference between b and a; that is, a − b.

[Runtime Function] _Decimal32 __dpd_mulsd3 (Decimal32 a, Decimal32 b)  
[Runtime Function] _Decimal32 __bid_mulsd3 (Decimal32 a, Decimal32 b)  
[Runtime Function] _Decimal64 __dpd_muldd3 (Decimal64 a, Decimal64 b)  
[Runtime Function] _Decimal64 __bid_muldd3 (Decimal64 a, Decimal64 b)  
[Runtime Function] _Decimal128 __dpd_multd3 (Decimal128 a, Decimal128 b)  
[Runtime Function] _Decimal128 __bid_multd3 (Decimal128 a, Decimal128 b)  

These functions return the product of a and b.

[Runtime Function] _Decimal32 __dpd_divsd3 (Decimal32 a, Decimal32 b)  
[Runtime Function] _Decimal32 __bid_divsd3 (Decimal32 a, Decimal32 b)  
[Runtime Function] _Decimal64 __dpd_divdd3 (Decimal64 a, Decimal64 b)  
[Runtime Function] _Decimal64 __bid_divdd3 (Decimal64 a, Decimal64 b)  
[Runtime Function] _Decimal128 __dpd_divtd3 (Decimal128 a, Decimal128 b)  
[Runtime Function] _Decimal128 __bid_divtd3 (Decimal128 a, Decimal128 b)  

These functions return the quotient of a and b; that is, a/b.

[Runtime Function] _Decimal32 __dpd_negsd2 (Decimal32 a)  
[Runtime Function] _Decimal32 __bid_negsd2 (Decimal32 a)  
[Runtime Function] _Decimal64 __dpd_negdd2 (Decimal64 a)  
[Runtime Function] _Decimal64 __bid_negdd2 (Decimal64 a)  
[Runtime Function] _Decimal128 __dpd_negtd2 (Decimal128 a)  
[Runtime Function] _Decimal128 __bid_negtd2 (Decimal128 a)  

These functions return the negation of a. They simply flip the sign bit, so they can produce negative zero and negative NaN.

4.3.2 Conversion functions

[Runtime Function] _Decimal64 __dpd_extendsddd2 (Decimal32 a)  
[Runtime Function] _Decimal64 __bid_extendsddd2 (Decimal32 a)  
[Runtime Function] _Decimal128 __dpd_extendsdtd2 (Decimal32 a)  
[Runtime Function] _Decimal128 __bid_extendsdtd2 (Decimal32 a)
These functions convert the value of \( a \) from one decimal floating type to another.

```c
_decimal64 __dpd_extendddtd2 (_decimal64 a) [Runtime Function]
_decimal64 __bid_extendddtd2 (_decimal64 a) [Runtime Function]
_decimal32 __dpd_truncddsd2 (_decimal64 a) [Runtime Function]
_decimal32 __bid_truncddsd2 (_decimal64 a) [Runtime Function]
_decimal32 __dpd_trunctdsd2 (_decimal128 a) [Runtime Function]
_decimal32 __bid_trunctdsd2 (_decimal128 a) [Runtime Function]
_decimal64 __dpd_trunctddd2 (_decimal128 a) [Runtime Function]
_decimal64 __bid_trunctddd2 (_decimal128 a) [Runtime Function]
```

These functions convert the value of \( a \) from a binary floating type to a decimal floating type of a different size.

```c
_decimal64 __dpd_extendsfdd (float a) [Runtime Function]
_decimal64 __bid_extendsfdd (float a) [Runtime Function]
_decimal128 __dpd_extendsftd (float a) [Runtime Function]
_decimal128 __bid_extendsftd (float a) [Runtime Function]
_decimal128 __dpd_trunctdsf (_decimal32 a) [Runtime Function]
_decimal128 __bid_trunctdsf (_decimal32 a) [Runtime Function]
_decimal64 __dpd_trunctddf (_decimal128 a) [Runtime Function]
_decimal64 __bid_trunctddf (_decimal128 a) [Runtime Function]
_decimal64 __dpd_extendsdf (_decimal32 a) [Runtime Function]
_decimal64 __bid_extendsdf (_decimal32 a) [Runtime Function]
_decimal64 __dpd_trunctdf (_decimal128 a) [Runtime Function]
_decimal64 __bid_trunctdf (_decimal128 a) [Runtime Function]
```

```c
float __dpd_trunccdsf (_decimal64 a) [Runtime Function]
float __bid_trunccdsf (_decimal64 a) [Runtime Function]
float __dpd_trunctdsf (_decimal128 a) [Runtime Function]
float __bid_trunctdsf (_decimal128 a) [Runtime Function]
double __dpd_extendsddf (_decimal32 a) [Runtime Function]
double __bid_extendsddf (_decimal32 a) [Runtime Function]
double __dpd_trunctddf (_decimal128 a) [Runtime Function]
double __bid_trunctddf (_decimal128 a) [Runtime Function]
long double __dpd_extendsdf (_decimal32 a) [Runtime Function]
long double __bid_extendsdf (_decimal32 a) [Runtime Function]
long double __dpd_trunctdf (_decimal128 a) [Runtime Function]
long double __bid_trunctdf (_decimal128 a) [Runtime Function]
long double __dpd_extendsddf (_decimal32 a) [Runtime Function]
long double __bid_extendsddf (_decimal32 a) [Runtime Function]
long double __dpd_extendsdf (_decimal32 a) [Runtime Function]
long double __bid_extendsdf (_decimal32 a) [Runtime Function]
```
these functions convert the value of \( a \) from a decimal floating type to a binary floating type of a different size.

These functions convert the value of \( a \) between decimal and binary floating types of the same size.

These functions convert \( a \) to a signed integer.

These functions convert \( a \) to a signed long.

These functions convert \( a \) to an unsigned integer. Negative values all become zero.
unsigned long __bid_fixunstddi (.Decimal128 a) [Runtime Function]

These functions convert a to an unsigned long. Negative values all become zero.

_DECIMAL32 __dpd_floatsisd (int i) [Runtime Function]
_DECIMAL32 __bid_floatsisd (int i) [Runtime Function]
_DECIMAL64 __dpd_floatsidd (int i) [Runtime Function]
_DECIMAL64 __bid_floatsidd (int i) [Runtime Function]
_DECIMAL128 __dpd_floatsitd (int i) [Runtime Function]
_DECIMAL128 __bid_floatsitd (int i) [Runtime Function]

These functions convert i, a signed integer, to decimal floating point.

_DECIMAL32 __dpd_floatdisd (long i) [Runtime Function]
_DECIMAL32 __bid_floatdisd (long i) [Runtime Function]
_DECIMAL64 __dpd_floatdidd (long i) [Runtime Function]
_DECIMAL64 __bid_floatdidd (long i) [Runtime Function]
_DECIMAL128 __dpd_floatditd (long i) [Runtime Function]
_DECIMAL128 __bid_floatditd (long i) [Runtime Function]

These functions convert i, a signed long, to decimal floating point.

_DECIMAL32 __dpd_floatunssisd (unsigned int i) [Runtime Function]
_DECIMAL32 __bid_floatunssisd (unsigned int i) [Runtime Function]
_DECIMAL64 __dpd_floatunssidd (unsigned int i) [Runtime Function]
_DECIMAL64 __bid_floatunssidd (unsigned int i) [Runtime Function]
_DECIMAL128 __dpd_floatunssitd (unsigned int i) [Runtime Function]
_DECIMAL128 __bid_floatunssitd (unsigned int i) [Runtime Function]

These functions convert i, an unsigned integer, to decimal floating point.

_DECIMAL32 __dpd_floatunssdisd (unsigned long i) [Runtime Function]
_DECIMAL32 __bid_floatunssdisd (unsigned long i) [Runtime Function]
_DECIMAL64 __dpd_floatunssdidd (unsigned long i) [Runtime Function]
_DECIMAL64 __bid_floatunssdidd (unsigned long i) [Runtime Function]
_DECIMAL128 __dpd_floatunssditd (unsigned long i) [Runtime Function]
_DECIMAL128 __bid_floatunssditd (unsigned long i) [Runtime Function]

These functions convert i, an unsigned long, to decimal floating point.

4.3.3 Comparison functions

int __dpd_unordsd2 (.Decimal32 a, .Decimal32 b) [Runtime Function]
int __bid_unordsd2 (.Decimal32 a, .Decimal32 b) [Runtime Function]
int __dpd_unorddd2 (.Decimal64 a, .Decimal64 b) [Runtime Function]
int __bid_unorddd2 (.Decimal64 a, .Decimal64 b) [Runtime Function]
int __dpd_unordtd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]
int __bid_unordtd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]

These functions return a nonzero value if either argument is NaN, otherwise 0.

There is also a complete group of higher level functions which correspond directly to
comparison operators. They implement the ISO C semantics for floating-point comparisons,
taking NaN into account. Pay careful attention to the return values defined for each set.
Under the hood, all of these routines are implemented as
if (__bid_unorder2 (a, b))
    return E;
return __bid_cmp2 (a, b);

where E is a constant chosen to give the proper behavior for NaN. Thus, the meaning of the return value is different for each set. Do not rely on this implementation; only the semantics documented below are guaranteed.

int __dpd_eqsd2 (.Decimal32 a, .Decimal32 b)  [Runtime Function]
int __bid_eqsd2 (.Decimal32 a, .Decimal32 b)  [Runtime Function]
int __dpd_eqdd2 (.Decimal64 a, .Decimal64 b)  [Runtime Function]
int __bid_eqdd2 (.Decimal64 a, .Decimal64 b)  [Runtime Function]
int __dpd_eqtd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]
int __bid_eqtd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]

These functions return zero if neither argument is NaN, and a and b are equal.

int __dpd_nesd2 (.Decimal32 a, .Decimal32 b)  [Runtime Function]
int __bid_nesd2 (.Decimal32 a, .Decimal32 b)  [Runtime Function]
int __dpd_nedd2 (.Decimal64 a, .Decimal64 b)  [Runtime Function]
int __bid_nedd2 (.Decimal64 a, .Decimal64 b)  [Runtime Function]
int __dpd_netd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]
int __bid_netd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]

These functions return a nonzero value if either argument is NaN, or if a and b are unequal.

int __dpd_gesd2 (.Decimal32 a, .Decimal32 b)  [Runtime Function]
int __bid_gesd2 (.Decimal32 a, .Decimal32 b)  [Runtime Function]
int __dpd_gedd2 (.Decimal64 a, .Decimal64 b)  [Runtime Function]
int __bid_gedd2 (.Decimal64 a, .Decimal64 b)  [Runtime Function]
int __dpd_gttd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]
int __bid_gttd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]

These functions return a value greater than or equal to zero if neither argument is NaN, and a is greater than or equal to b.

int __dpd_ltsd2 (.Decimal32 a, .Decimal32 b)  [Runtime Function]
int __bid_ltsd2 (.Decimal32 a, .Decimal32 b)  [Runtime Function]
int __dpd_ltd2 (.Decimal64 a, .Decimal64 b)  [Runtime Function]
int __bid_ltd2 (.Decimal64 a, .Decimal64 b)  [Runtime Function]
int __dpd_lttd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]
int __bid_lttd2 (.Decimal128 a, .Decimal128 b) [Runtime Function]

These functions return a value less than zero if neither argument is NaN, and a is strictly less than b.
int __bid_letd2 (Decimal128 a, Decimal128 b)  
These functions return a value less than or equal to zero if neither argument is NaN, 
and a is less than or equal to b.

int __dpd_gtsd2 (Decimal32 a, Decimal32 b)  
int __bid_gtsd2 (Decimal32 a, Decimal32 b)  
int __dpd_gtdd2 (Decimal64 a, Decimal64 b)  
int __bid_gtdd2 (Decimal64 a, Decimal64 b)  
int __dpd_gttd2 (Decimal128 a, Decimal128 b)  
int __bid_gttd2 (Decimal128 a, Decimal128 b)  
These functions return a value greater than zero if neither argument is NaN, and 
a is strictly greater than b.

4.4 Routines for fixed-point fractional emulation

The software fixed-point library implements fixed-point fractional arithmetic, and is only 
activated on selected targets.

For ease of comprehension fract is an alias for the _Fract type, accum an alias for 
_Accum, and sat an alias for _Sat.

For illustrative purposes, in this section the fixed-point fractional type short fract is as-
sumed to correspond to machine mode QQmode; unsigned short fract to UQQmode; fract 
to HQmode; unsigned fract to UHQmode; long fract to SQmode; unsigned long fract 
to USQmode; long long fract to DQmode; and unsigned long long fract to UDQmode. 
Similarly the fixed-point accumulator type short accum corresponds to HAmode; 
unsigned short accum to UHAMode; accum to SAmode; unsigned accum to USAmode; 
long accum to DAmode; unsigned long accum to UDAmode; long long accum to TAmode; 
and unsigned long long accum to UTAmode.

4.4.1 Arithmetic functions

short fract __addqq3 (short fract a, short fract b)  
fract __addhq3 (fract a, fract b)  
long fract __addsq3 (long fract a, long fract b)  
long long fract __adddq3 (long long fract a, long long fract 
b)  
unsigned short fract __adduqq3 (unsigned short fract a, 
unsigned short fract b)  
unsigned fract __adduhq3 (unsigned fract a, unsigned fract 
b)  
unsigned long fract __addusq3 (unsigned long fract a, 
unsigned long fract b)  
unsigned long long fract __addudq3 (unsigned long long 
fract a, unsigned long long fract b)  
short accum __addha3 (short accum a, short accum b)  
accum __addsa3 (accum a, accum b)  
long accum __addda3 (long accum a, long accum b)  
long long accum __addta3 (long long accum a, long long 
accum b)
These functions return the sum of \( a \) and \( b \).

unsigned short accum __adduha3 (unsigned short accum a, unsigned short accum b)  
unsigned accum __addusa3 (unsigned accum a, unsigned accum b)  
unsigned long accum __adduda3 (unsigned long accum a, unsigned long accum b)  
unsigned long long accum __adduta3 (unsigned long long accum a, unsigned long long accum b)

These functions return the sum of \( a \) and \( b \) with signed saturation.

short fract __ssaddqq3 (short fract a, short fract b)  
fract __ssaddhq3 (fract a, fract b)  
long fract __ssaddsq3 (long fract a, long fract b)  
long long fract __ssadddq3 (long long fract a, long long fract b)

These functions return the sum of \( a \) and \( b \) with unsigned saturation.

unsigned short fract __usadduqq3 (unsigned short fract a, unsigned short fract b)  
unsigned fract __usadduhq3 (unsigned fract a, unsigned fract b)  
unsigned long fract __usaddusq3 (unsigned long fract a, unsigned long fract b)  
unsigned long long fract __usaddudq3 (unsigned long long fract a, unsigned long long fract b)

These functions return the sum of \( a \) and \( b \).
unsigned fract __subuhq3 (unsigned fract a, unsigned fract b)  
unsigned long fract __subusq3 (unsigned long fract a, unsigned long fract b)  
unsigned long long fract __subudq3 (unsigned long long fract a, unsigned long long fract b)  
short accum __subha3 (short accum a, short accum b)  
accum __subsa3 (accum a, accum b)  
long accum __subda3 (long accum a, long accum b)  
long long accum __subta3 (long long accum a, long long accum b)  
unsigned short accum __subuhq3 (unsigned short accum a, unsigned short accum b)  
unsigned accum __subusa3 (unsigned accum a, unsigned accum b)  
unsigned long accum __subuda3 (unsigned long accum a, unsigned long accum b)  
unsigned long long accum __subuta3 (unsigned long long accum a, unsigned long long accum b)  

These functions return the difference of a and b; that is, a - b.

short fract __sssubqq3 (short fract a, short fract b)  
fract __sssubhq3 (fract a, fract b)  
long fract __sssubsq3 (long fract a, long fract b)  
long long fract __sssubdq3 (long long fract a, long long fract b)  
short accum __sssubha3 (short accum a, short accum b)  
accum __sssubsa3 (accum a, accum b)  
long accum __sssubda3 (long accum a, long accum b)  
long long accum __sssubta3 (long long accum a, long long accum b)  

These functions return the difference of a and b with signed saturation; that is, a - b.
unsigned long long accum __ussubuta3 (unsigned long
long accum a, unsigned long long accum b)

These functions return the difference of a and b with unsigned saturation; that is, \(a - b\).

short fract __mulqq3 (short fract a, short fract b)
fract __mulhq3 (fract a, fract b)
long fract __mulsq3 (long fract a, long fract b)
long long fract __muldq3 (long long fract a, long long fract b)
unsigned short fract __muluqq3 (unsigned short fract a,
unsigned short fract b)
unsigned fract __muluhq3 (unsigned fract a, unsigned fract b)
unsigned long fract __mulusq3 (unsigned long fract a,
unsigned long fract b)
unsigned long long fract __muludq3 (unsigned long long fract a, unsigned long long fract b)
short accum __mulha3 (short accum a, short accum b)
accum __mulsa3 (accum a, accum b)
long accum __mulda3 (long accum a, long accum b)
long long accum __multa3 (long long accum a, long long accum b)
unsigned short accum __muluha3 (unsigned short accum a,
unsigned short accum b)
unsigned accum __mulusa3 (unsigned accum a, unsigned accum b)
unsigned long accum __muluda3 (unsigned long accum a,
unsigned long accum b)
unsigned long long accum __muluta3 (unsigned long long accum a, unsigned long long accum b)

These functions return the product of a and b with signed saturation.
unsigned short fract __usmuluqq3 (unsigned short fract a, unsigned short fract b) [Runtime Function]
unsigned fract __usmuluhq3 (unsigned fract a, unsigned fract b) [Runtime Function]
unsigned long fract __usmulusq3 (unsigned long fract a, unsigned long fract b) [Runtime Function]
unsigned long long fract __usmuludq3 (unsigned long long fract a, unsigned long long fract b) [Runtime Function]
unsigned short accum __usmuluha3 (unsigned short accum a, unsigned short accum b) [Runtime Function]
unsigned accum __usmulusa3 (unsigned accum a, unsigned accum b) [Runtime Function]
unsigned long accum __usmuluda3 (unsigned long accum a, unsigned long accum b) [Runtime Function]
unsigned long long accum __usmuluta3 (unsigned long long accum a, unsigned long long accum b) [Runtime Function]

These functions return the product of a and b with unsigned saturation.

short fract __divqq3 (short fract a, short fract b) [Runtime Function]
fract __divhq3 (fract a, fract b) [Runtime Function]
long fract __divsq3 (long fract a, long fract b) [Runtime Function]
long long fract __divdq3 (long long fract a, long long fract b) [Runtime Function]
short accum __divha3 (short accum a, short accum b) [Runtime Function]
accum __divsa3 (accum a, accum b) [Runtime Function]
long accum __divda3 (long accum a, long accum b) [Runtime Function]
long long accum __divta3 (long long accum a, long long accum b) [Runtime Function]

These functions return the quotient of the signed division of a and b.

unsigned short fract __udivuqq3 (unsigned short fract a, unsigned short fract b) [Runtime Function]
unsigned fract __udivuhq3 (unsigned fract a, unsigned fract b) [Runtime Function]
unsigned long fract __udivusq3 (unsigned long fract a, unsigned long fract b) [Runtime Function]
unsigned long long fract __udivudq3 (unsigned long long fract a, unsigned long long fract b) [Runtime Function]
unsigned short accum __udivuha3 (unsigned short accum a, unsigned short accum b) [Runtime Function]
unsigned accum __udivusa3 (unsigned accum a, unsigned accum b) [Runtime Function]
unsigned long accum __udivuda3 (unsigned long accum a, unsigned long accum b) [Runtime Function]
unsigned long long accum __udivuta3 (unsigned long long accum a, unsigned long long accum b) [Runtime Function]

These functions return the quotient of the unsigned division of a and b.
short fract __ssdivqq3 (short fract a, short fract b)
fract __ssdivhq3 (fract a, fract b)
long fract __ssdivsq3 (long fract a, long fract b)
long long fract __ssdivdq3 (long long fract a, long long fract b)

short accum __ssdivha3 (short accum a, short accum b)
accum __ssdivsa3 (accum a, accum b)
long accum __ssdivda3 (long accum a, long accum b)
long long accum __ssdivta3 (long long accum a, long long accum b)

These functions return the quotient of the signed division of a and b with signed saturation.

unsigned short fract __usdivuqq3 (unsigned short fract a, unsigned short fract b)
unsigned fract __usdivuhq3 (unsigned fract a, unsigned fract b)
unsigned long fract __usdivusq3 (unsigned long fract a, unsigned long fract b)
unsigned long long fract __usdivudq3 (unsigned long long fract a, unsigned long long fract b)
unsigned short accum __usdivuha3 (unsigned short accum a, unsigned short accum b)
unsigned accum __usdivusa3 (unsigned accum a, unsigned accum b)
unsigned long accum __usdivuda3 (unsigned long accum a, unsigned long accum b)
unsigned long long accum __usdivuta3 (unsigned long long accum a, unsigned long long accum b)

These functions return the quotient of the unsigned division of a and b with unsigned saturation.
unsigned long long accum __neguta2 (unsigned long long accum a)  

These functions return the negation of a.

short fract __ssnegqq2 (short fract a)  
fract __ssneghq2 (fract a)  
long fract __ssnegsq2 (long fract a)  
long long fract __ssnegdq2 (long long fract a)  
short accum __ssnegha2 (short accum a)  
accum __ssnegsa2 (accum a)  
long accum __ssnegda2 (long accum a)  
long long accum __ssnegta2 (long long accum a)  

These functions return the negation of a with signed saturation.

unsigned short fract __usneguqq2 (unsigned short fract a)  
unsigned fract __usneguhq2 (unsigned fract a)  
unsigned long fract __usnegusq2 (unsigned long fract a)  
unsigned long long fract __usnegudq2 (unsigned long long fract a)  
unsigned short accum __usneguha2 (unsigned short accum a)  
unsigned accum __usnegusa2 (unsigned accum a)  
unsigned long accum __usneguda2 (unsigned long accum a)  
unsigned long long accum __usneguta2 (unsigned long long accum a)  

These functions return the negation of a with unsigned saturation.
unsigned long long accum __ashluta3 (unsigned long long accum a, int b)
These functions return the result of shifting a left by b bits.

short fract __ashrq3 (short fract a, int b)
fract __ashrhq3 (fract a, int b)
long fract __ashrsq3 (long fract a, int b)
long long fract __ashrdq3 (long long fract a, int b)
short accum __ashrha3 (short accum a, int b)
accum __ashrsqa3 (accum a, int b)
long accum __ashrda3 (long accum a, int b)
long long accum __ashrta3 (long long accum a, int b)
These functions return the result of arithmetically shifting a right by b bits.

unsigned short fract __lshruqq3 (unsigned short fract a, int b)
unsigned fract __lshruhq3 (unsigned fract a, int b)
unsigned long fract __lshrusq3 (unsigned long fract a, int b)
unsigned long long fract __lshrudq3 (unsigned long long fract a, int b)
unsigned short accum __lshruha3 (unsigned short accum a, int b)
unsigned accum __lshrusa3 (unsigned accum a, int b)
unsigned long accum __lshruda3 (unsigned long accum a, int b)
unsigned long long accum __lshruta3 (unsigned long long accum a, int b)
These functions return the result of logically shifting a right by b bits.

fract __ssashlhq3 (fract a, int b)
long fract __ssashlsq3 (long fract a, int b)
long long fract __ssashldq3 (long long fract a, int b)
short accum __ssashlha3 (short accum a, int b)
accum __ssashlsa3 (accum a, int b)
long accum __ssashlda3 (long accum a, int b)
long long accum __ssashlta3 (long long accum a, int b)
These functions return the result of shifting a left by b bits with signed saturation.

unsigned short fract __usashluqq3 (unsigned short fract a, int b)
unsigned fract __usashluhq3 (unsigned fract a, int b)
unsigned long fract __usashlusq3 (unsigned long fract a, int b)
unsigned long long fract __usashludq3 (unsigned long long fract a, int b)
unsigned short accum __usashluha3 (unsigned short accum a, int b)

unsigned accum __usashlusa3 (unsigned accum a, int b) [Runtime Function]
unsigned long accum __usashluda3 (unsigned long accum a, int b)
unsigned long long accum __usashluta3 (unsigned long long accum a, int b)

These functions return the result of shifting a left by b bits with unsigned saturation.

4.4.2 Comparison functions

The following functions implement fixed-point comparisons. These functions implement a low-level compare, upon which the higher level comparison operators (such as less than and greater than or equal to) can be constructed. The returned values lie in the range zero to two, to allow the high-level operators to be implemented by testing the returned result using either signed or unsigned comparison.

int __cmpqq2 (short fract a, short fract b) [Runtime Function]
int __cmphq2 (fract a, fract b) [Runtime Function]
int __cmphq2 (fract a, fract b) [Runtime Function]
int __cmpsq2 (long fract a, long fract b) [Runtime Function]
int __cmpdq2 (long long fract a, long long fract b) [Runtime Function]
int __cmpuqq2 (unsigned short fract a, unsigned short fract b) [Runtime Function]
int __cmpuhq2 (unsigned fract a, unsigned fract b) [Runtime Function]
int __cmpusq2 (unsigned long fract a, unsigned long fract b) [Runtime Function]
int __cmpudq2 (unsigned long long fract a, unsigned long long fract b) [Runtime Function]
int __cmpha2 (short accum a, short accum b) [Runtime Function]
int __cmpsa2 (accum a, accum b) [Runtime Function]
int __cmpha2 (short accum a, short accum b) [Runtime Function]
int __cmpda2 (long accum a, long accum b) [Runtime Function]
int __cmpta2 (long long accum a, long long accum b) [Runtime Function]
int __cmpha2 (short accum a, short accum b) [Runtime Function]
int __cmpusa2 (unsigned accum a, unsigned accum b) [Runtime Function]
int __cmpuda2 (unsigned long accum a, unsigned long accum b) [Runtime Function]
int __cmputa2 (unsigned long long accum a, unsigned long long accum b) [Runtime Function]

These functions perform a signed or unsigned comparison of a and b (depending on the selected machine mode). If a is less than b, they return 0; if a is greater than b, they return 2; and if a and b are equal they return 1.

4.4.3 Conversion functions

fract __fractqqh2 (short fract a) [Runtime Function]
long fract __fractqqsq2 (short fract a) [Runtime Function]
long long fract __fractqqdq2 (short fract a) [Runtime Function]
short accum __fractqha (short fract a) [Runtime Function]
accum __fractqqsa (short fract a) [Runtime Function]
long accum __fractqqda (short fract a) [Runtime Function]
long long accum __fractqqta (short fract a) [Runtime Function]
unsigned short fract __fractqqquqq (short fract a) [Runtime Function]
unsigned fract __fractqquhq (short fract a) [Runtime Function]
unsigned long fract __fractqusq (short fract a)  [Runtime Function]
unsigned long long fract __fractqqudq (short fract a)  [Runtime Function]
unsigned short accum __fractqquha (short fract a)  [Runtime Function]
unsigned accum __fractqqusa (short fract a)  [Runtime Function]
unsigned long accum __fractqquda (short fract a)  [Runtime Function]
unsigned long long accum __fractqquta (short fract a)  [Runtime Function]
signed char __fractqqqi (short fract a)  [Runtime Function]
short __fractqqhi (short fract a)  [Runtime Function]
int __fractqqsi (short fract a)  [Runtime Function]
long __fractqqdi (short fract a)  [Runtime Function]
long long __fractqqti (short fract a)  [Runtime Function]
float __fractqqsf (short fract a)  [Runtime Function]
double __fractqqdf (short fract a)  [Runtime Function]
short fract __fracthqqq2 (fract a)  [Runtime Function]
long fract __fracthqsq2 (fract a)  [Runtime Function]
long long fract __fracthqdq2 (fract a)  [Runtime Function]
short accum __fracthqha (fract a)  [Runtime Function]
accum __fracthqsa (fract a)  [Runtime Function]
long accum __fracthqda (fract a)  [Runtime Function]
long long accum __fracthqta (fract a)  [Runtime Function]
unsigned short fract __fracthuqq (fract a)  [Runtime Function]
unsigned fract __fracthuhq (fract a)  [Runtime Function]
unsigned long fract __fracthuqsq (fract a)  [Runtime Function]
unsigned long long fract __fracthqudq (fract a)  [Runtime Function]
unsigned short accum __fracthuqa (fract a)  [Runtime Function]
unsigned accum __fracthuqa (fract a)  [Runtime Function]
unsigned long accum __fracthquda (fract a)  [Runtime Function]
unsigned long long accum __fracthquta (fract a)  [Runtime Function]
signed char __fracthqqi (fract a)  [Runtime Function]
short __fracthqhi (fract a)  [Runtime Function]
int __fracthqsi (fract a)  [Runtime Function]
long __fracthqi (fract a)  [Runtime Function]
long long __fracthqi (fract a)  [Runtime Function]
float __fracthqsf (fract a)  [Runtime Function]
double __fracthqdf (fract a)  [Runtime Function]
short fract __fractsqqq2 (long fract a)  [Runtime Function]
fract __fractsqhq2 (long fract a)  [Runtime Function]
long long fract __fractsqsdq2 (long fract a)  [Runtime Function]
short accum __fractsqha (long fract a)  [Runtime Function]
accum __fractsqsa (long fract a)  [Runtime Function]
long accum __fractsqda (long fract a)  [Runtime Function]
long long accum __fractsqta (long fract a)  [Runtime Function]
unsigned short fract __fractsqquq (long fract a)  [Runtime Function]
unsigned fract __fractsquhq (long fract a)  [Runtime Function]
unsigned long fract __fractsquqsq (long fract a)  [Runtime Function]
unsigned long long fract __fractsquqdm (long fract a)  [Runtime Function]
unsigned short accum __fractsqquha (long fract a)  [Runtime Function]
unsigned accum __fractsqusa (long fract a)  
unsigned long accum __fractsquda (long fract a)  
unsigned long long accum __fractsquta (long fract a)  
signed char __fractsqqi (long fract a)  
short __fractsqhi (long fract a)  
int __fractsqsi (long fract a)  
long __fractsqdi (long fract a)  
long long __fractsqti (long fract a)  
float __fractsqsf (long fract a)  
double __fractsqdf (long fract a)  
short fract __fractdqq2 (long long fract a)  
fract __fractdqhq2 (long long fract a)  
long fract __fractdqsq2 (long long fract a)  
short accum __fractdqha (long long fract a)  
accum __fractdqua (long long fract a)  
long accum __fractdquda (long long fract a)  
long long accum __fractdquta (long long fract a)  
unsigned short fract __fractdqqq (long long fract a)  
unsigned fract __fractdqhq (long long fract a)  
unsigned long fract __fractdqsq (long long fract a)  
unsigned long long fract __fractdqud (long long fract a)  
unsigned short accum __fractdqha (long long fract a)  
unsigned accum __fractdqusa (long long fract a)  
unsigned long accum __fractdquda (long long fract a)  
unsigned long long accum __fractdquta (long long fract a)  
signed char __fractdqqi (long long fract a)  
short __fractdqi (long long fract a)  
int __fractdqi (long long fract a)  
long __fractdqi (long long fract a)  
long long __fractdqi (long long fract a)  
float __fractdqi (long long fract a)  
double __fractdqi (long long fract a)  
short fract __fractdquq (short accum a)  
fract __fractdquhq (short accum a)  
long fract __fractdqsq (short accum a)  
long long fract __fractdqud (short accum a)  
accum __fractdusa (short accum a)  
long accum __fractduda (short accum a)  
long long accum __fractduda2 (short accum a)  
unsigned short fract __fractdauqq (short accum a)  
unsigned fract __fractdauhq (short accum a)  
unsigned long fract __fractdusq (short accum a)  
unsigned long long fract __fractdauuq (short accum a)  
unsigned short accum __fractdauha (short accum a)  
unsigned accum __fractdusqa (short accum a)
unsigned long accum __fracthauda (short accum a) [Runtime Function]
unsigned long long accum __fracthauta (short accum a) [Runtime Function]
signed char __fracthaqi (short accum a) [Runtime Function]
short __fracthahi (short accum a) [Runtime Function]
int __fracthasi (short accum a) [Runtime Function]
long __fracthadi (short accum a) [Runtime Function]
long long __fracthadi (short accum a) [Runtime Function]
float __fracthasf (short accum a) [Runtime Function]
double __fracthadf (short accum a) [Runtime Function]
short fract __fractsaqq (accum a) [Runtime Function]
fract __fractsahq (accum a) [Runtime Function]
long fract __fractsaqq (accum a) [Runtime Function]
long long fract __fractsaqq (accum a) [Runtime Function]
short accum __fractsaha2 (accum a) [Runtime Function]
long accum __fractsada2 (accum a) [Runtime Function]
long long accum __fractsada2 (accum a) [Runtime Function]
unsigned short fract __fractsaauqq (accum a) [Runtime Function]
unsigned fract __fractsaauhq (accum a) [Runtime Function]
unsigned long fract __fractsausq (accum a) [Runtime Function]
unsigned long long fract __fractsausq (accum a) [Runtime Function]
unsigned short accum __fractsauha (accum a) [Runtime Function]
unsigned accum __fractsausa (accum a) [Runtime Function]
unsigned long accum __fractsauda (accum a) [Runtime Function]
unsigned long long accum __fractsauda (accum a) [Runtime Function]
signed char __fractsaqi (accum a) [Runtime Function]
short __fractsahi (accum a) [Runtime Function]
int __fractsasi (accum a) [Runtime Function]
long __fractsadi (accum a) [Runtime Function]
long long __fractsati (accum a) [Runtime Function]
float __fractsasf (accum a) [Runtime Function]
double __fractsadf (accum a) [Runtime Function]
short fract __fractddaqq (long accum a) [Runtime Function]
fract __fractdahq (long accum a) [Runtime Function]
long fract __fractdasaq (long accum a) [Runtime Function]
long long fract __fractdasaq (long accum a) [Runtime Function]
short accum __fractdaha2 (long accum a) [Runtime Function]
accum __fractdasa2 (long accum a) [Runtime Function]
long long accum __fractdata2 (long accum a) [Runtime Function]
unsigned short fract __fractdauqq (long accum a) [Runtime Function]
unsigned fract __fractdauhq (long accum a) [Runtime Function]
unsigned long fract __fractdausq (long accum a) [Runtime Function]
unsigned long long fract __fractdausq (long accum a) [Runtime Function]
unsigned short accum __fractdauha (long accum a) [Runtime Function]
unsigned accum __fractdausa (long accum a) [Runtime Function]
unsigned long accum __fractdaua (long accum a) [Runtime Function]
unsigned long long accum __fractdaua (long accum a) [Runtime Function]
signed char __fractdaqi (long accum a) [Runtime Function]
short __fractdahi (long accum a) [Runtime Function]
int __fractdasi (long accum a) [Runtime Function]
long __fractdadi (long accum a) [Runtime Function]
long long __fractdati (long accum a) [Runtime Function]
float __fractdasf (long accum a) [Runtime Function]
double __fractdadf (long accum a) [Runtime Function]
short fract __fracttaqq (long long accum a) [Runtime Function]
fract __fracttahq (long long accum a) [Runtime Function]
long fract __fracttasq (long long accum a) [Runtime Function]
long long fract __fracttadq (long long accum a) [Runtime Function]
short accum __fracttaha2 (long long accum a) [Runtime Function]
accum __fracttasa2 (long long accum a) [Runtime Function]
long accum __fracttada2 (long long accum a) [Runtime Function]
unsigned short fract __fracttauqq (long long accum a) [Runtime Function]
unsigned fract __fracttauhq (long long accum a) [Runtime Function]
unsigned long fract __fracttausq (long long accum a) [Runtime Function]
unsigned long long fract __fracttaudq (long long accum a) [Runtime Function]
unsigned short accum __fracttahuha (long long accum a) [Runtime Function]
unsigned accum __fracttahuasa (long long accum a) [Runtime Function]
unsigned long accum __fracttahuada (long long accum a) [Runtime Function]
unsigned long long accum __fracttahuata (long long accum a) [Runtime Function]
signed char __fracttaqi (long long accum a) [Runtime Function]
short __fracttahi (long long accum a) [Runtime Function]
int __fracttasi (long long accum a) [Runtime Function]
long __fracttadi (long long accum a) [Runtime Function]
long long __fracttati (long long accum a) [Runtime Function]
float __fracttasf (long long accum a) [Runtime Function]
double __fracttadf (long long accum a) [Runtime Function]
short fract __fracttqqqq (unsigned short fract a) [Runtime Function]
fract __fracttuqqhq (unsigned short fract a) [Runtime Function]
long fract __fracttuqqsq (unsigned short fract a) [Runtime Function]
long long fract __fracttuqqdq (unsigned short fract a) [Runtime Function]
short accum __fracttuqqha (unsigned short fract a) [Runtime Function]
accum __fracttuqqsa (unsigned short fract a) [Runtime Function]
long accum __fracttuqqda (unsigned short fract a) [Runtime Function]
long long accum __fracttuqqta (unsigned short fract a) [Runtime Function]
unsigned fract __fracttuqqhq2 (unsigned short fract a) [Runtime Function]
unsigned long fract __fracttuqqusq2 (unsigned short fract a) [Runtime Function]
unsigned long long fract __fracttuqqudq2 (unsigned short fract a) [Runtime Function]
unsigned short accum __fracttuquha (unsigned short fract a) [Runtime Function]
unsigned accum __fracttuqusa (unsigned short fract a) [Runtime Function]
unsigned long accum__fractuqquadu(unsigned short fract a) [Runtime Function]
unsigned long long accum__fractuqquta(unsigned short fract a) [Runtime Function]
signed char __fractuqqqi(unsigned short fract a) [Runtime Function]
short __fractuqqhi(unsigned short fract a) [Runtime Function]
int __fractuqqsi(unsigned short fract a) [Runtime Function]
long __fractuqqdi(unsigned short fract a) [Runtime Function]
long long __fractuqqti(unsigned short fract a) [Runtime Function]
float __fractuqqsf(unsigned short fract a) [Runtime Function]
double __fractuqqdf(unsigned short fract a) [Runtime Function]
short fract __fractuhqqq(unsigned fract a) [Runtime Function]
fract __fractuhhqh(unsigned fract a) [Runtime Function]
long fract __fractuhhqq(unsigned fract a) [Runtime Function]
long long fract __fractuhqdq(unsigned fract a) [Runtime Function]
short accum__fractuhqha(unsigned fract a) [Runtime Function]
accum__fractuhqsa(unsigned fract a) [Runtime Function]
long accum__fractuhqda(unsigned fract a) [Runtime Function]
long long accum__fractuhqta(unsigned fract a) [Runtime Function]
unsigned short fract__fractusqqq(unsigned long fract a) [Runtime Function]
fract __fractusqhq(unsigned long fract a) [Runtime Function]
long fract __fractusqsq(unsigned long fract a) [Runtime Function]
short accum__fractusqha(unsigned long fract a) [Runtime Function]
accum__fractusqsa(unsigned long fract a) [Runtime Function]
long accum__fractusqda(unsigned long fract a) [Runtime Function]
signed char __fractusqqqi(unsigned long fract a) [Runtime Function]
short __fractusqhi(unsigned long fract a) [Runtime Function]
int __fractusqsi(unsigned long fract a) [Runtime Function]
long __fractusqdi(unsigned long fract a) [Runtime Function]
long long __fractusqti(unsigned long fract a) [Runtime Function]
float __fractusqsf(unsigned long fract a) [Runtime Function]
double __fractusqdf(unsigned long fract a) [Runtime Function]
unsigned short fract__fractusquqq2(unsigned long fract a) [Runtime Function]
unsigned fract__fractusquhq2(unsigned long fract a) [Runtime Function]
unsigned long fract__fractusqudq2(unsigned long fract a) [Runtime Function]
unsigned short fract__fractusquha(unsigned long fract a) [Runtime Function]
unsigned accum__fractusquha(unsigned long fract a) [Runtime Function]
unsigned long accum__fractusquta(unsigned long fract a) [Runtime Function]
unsigned long long fract __fractusqudq2 (unsigned long fract a)
unsigned short accum __fractusquha (unsigned long fract a)
unsigned accum __fractusqusa (unsigned long fract a)
unsigned long accum __fractusquda (unsigned long fract a)
unsigned long long accum __fractusquta (unsigned long fract a)
signed char __fractusqqi (unsigned long fract a)
short __fractusqqi (unsigned long fract a)
int __fractusqqi (unsigned long fract a)
long __fractusqqi (unsigned long fract a)
long long __fractusqqi (unsigned long fract a)
float __fractusqqf (unsigned long fract a)
double __fractusqqf (unsigned long fract a)
short fract __fractusqq2 (unsigned long long fract a)
unsigned __fractusqq2 (unsigned long long fract a)
unsigned long fract __fractusqq2 (unsigned long long fract a)
unsigned short accum __fractusqq2 (unsigned long long fract a)
unsigned long accum __fractusqq2 (unsigned long long fract a)
unsigned long long accum __fractusqq2 (unsigned long long fract a)
signed char __fractusqqi (unsigned long long fract a)
short __fractusqqi (unsigned long long fract a)
int __fractusqqi (unsigned long long fract a)
long __fractusqqi (unsigned long long fract a)
long long __fractusqqi (unsigned long long fract a)
float __fractusqqf (unsigned long long fract a)
double __fractusqqf (unsigned long long fract a)
short fract __fractusqq (unsigned long long fract a)
unsigned __fractusqq (unsigned long long fract a)
unsigned long fract __fractusqq (unsigned long long fract a)
unsigned short accum __fractusqq (unsigned long long fract a)
unsigned long accum __fractusqq (unsigned long long fract a)
unsigned long long accum __fractusqq (unsigned long long fract a)
short accum __fractuhaha (unsigned short accum a) [Runtime Function]
accum __fractuhasa (unsigned short accum a) [Runtime Function]
long accum __fractuhada (unsigned short accum a) [Runtime Function]
long long accum __fractuhata (unsigned short accum a) [Runtime Function]
unsigned short fract __fractuhaauqq (unsigned short
accum a) [Runtime Function]
unsigned fract __fractuhauhq (unsigned short accum a) [Runtime Function]
unsigned long fract __fractuhausq (unsigned short accum
a) [Runtime Function]
unsigned long long fract __fractuhaudq (unsigned short
accum a) [Runtime Function]
unsigned accum __fractuhausa2 (unsigned short accum a)
[Runtime Function]
unsigned long accum __fractuhauda2 (unsigned short accum
a) [Runtime Function]
unsigned long long accum __fractuhausta2 (unsigned short accum
a) [Runtime Function]
signed char __fractuhaqi (unsigned short accum a) [Runtime Function]
short __fractuhaqi (unsigned short accum a) [Runtime Function]
int __fractuhasi (unsigned short accum a) [Runtime Function]
long __fractuhadi (unsigned short accum a) [Runtime Function]
long long __fractuhati (unsigned short accum a) [Runtime Function]
float __fractuhasf (unsigned short accum a) [Runtime Function]
double __fractuhadf (unsigned short accum a) [Runtime Function]
short fract __fractusauqq (unsigned accum a) [Runtime Function]
fract __fractusahq (unsigned accum a) [Runtime Function]
long fract __fractusasq (unsigned accum a) [Runtime Function]
long long fract __fractusadq (unsigned accum a) [Runtime Function]
short accum __fractusaha (unsigned accum a) [Runtime Function]
accum __fractusasa (unsigned accum a) [Runtime Function]
long accum __fractusada (unsigned accum a) [Runtime Function]
long long accum __fractusata (unsigned accum a) [Runtime Function]
unsigned short fract __fractusauqq (unsigned accum a) [Runtime Function]
unsigned fract __fractusauhq (unsigned accum a) [Runtime Function]
unsigned long fract __fractusausq (unsigned accum a) [Runtime Function]
unsigned long long fract __fractusaudq (unsigned
accum a) [Runtime Function]
unsigned short accum __fractusauha2 (unsigned accum a)
[Runtime Function]
unsigned long accum __fractusauha2 (unsigned accum a)
[Runtime Function]
unsigned long long accum __fractusauhta2 (unsigned
accum a) [Runtime Function]
signed char __fractusaqi (unsigned accum a) [Runtime Function]
short __fractusahi (unsigned accum a) [Runtime Function]
int __fractusaqi (unsigned accum a) [Runtime Function]
long __fractusadi (unsigned accum a) [Runtime Function]
long long __fractusatdi (unsigned accum a) [Runtime Function]
float __fractusasf (unsigned accum a) [Runtime Function]
double __fractusadf (unsigned accum a) [Runtime Function]
short fract __fractudaqq (unsigned long accum a)  [Runtime Function]
fract __fractudahq (unsigned long accum a)  [Runtime Function]
long fract __fractudasq (unsigned long accum a)  [Runtime Function]
long long fract __fractudaq (unsigned long accum a)  [Runtime Function]
short accum __fractudaha (unsigned long accum a)  [Runtime Function]
accum __fractudasa (unsigned long accum a)  [Runtime Function]
long accum __fractudada (unsigned long accum a)  [Runtime Function]
long long accum __fractudata (unsigned long accum a)  [Runtime Function]
unsigned short fract __fractudauqq (unsigned long
accum a)  [Runtime Function]
unsigned fract __fractudauhq (unsigned long accum a)  [Runtime Function]
unsigned long fract __fractudausq (unsigned long accum a)  [Runtime Function]
unsigned long long fract __fractudaudq (unsigned long
accum a)  [Runtime Function]
unsigned short accum __fractudausha2 (unsigned long
accum a)  [Runtime Function]
unsigned accum __fractudausa2 (unsigned long accum a)  [Runtime Function]
unsigned long long accum __fractudauta2 (unsigned long
accum a)  [Runtime Function]
signed char __fractudaqi (unsigned long accum a)  [Runtime Function]
short __fractudah (unsigned long accum a)  [Runtime Function]
int __fractudas (unsigned long accum a)  [Runtime Function]
long __fractudadi (unsigned long accum a)  [Runtime Function]
long long __fractudati (unsigned long accum a)  [Runtime Function]
float __fractudasf (unsigned long accum a)  [Runtime Function]
double __fractudadf (unsigned long accum a)  [Runtime Function]
short fract __fractutaqq (unsigned long long accum a)  [Runtime Function]
fract __fractutahq (unsigned long long accum a)  [Runtime Function]
long fract __fractutasq (unsigned long long accum a)  [Runtime Function]
long long fract __fractutadq (unsigned long long accum a)  [Runtime Function]
short accum __fractutaha (unsigned long long accum a)  [Runtime Function]
accum __fractutasa (unsigned long long accum a)  [Runtime Function]
long accum __fractutada (unsigned long long accum a)  [Runtime Function]
long long accum __fractutata (unsigned long long accum a)  [Runtime Function]
unsigned short fract __fractutauqq (unsigned long long
accum a)  [Runtime Function]
unsigned fract __fractutauhq (unsigned long long accum a)  [Runtime Function]
unsigned long fract __fractutausq (unsigned long long
accum a)  [Runtime Function]
unsigned long long fract __fractutaudq (unsigned long long
accum a)  [Runtime Function]
unsigned short accum __fractutauha2 (unsigned long
accum a)  [Runtime Function]
unsigned accum __fractutausa2 (unsigned long long accum a)
unsigned long accum __fractutauda2 (unsigned long long accum a)
signed char __fractutaqi (unsigned long long accum a)
short __fractutahi (unsigned long long accum a)
int __fractutasi (unsigned long long accum a)
long __fractutadi (unsigned long long accum a)
long long __fractutati (unsigned long long accum a)
float __fractutasf (unsigned long long accum a)
double __fractutadfs (unsigned long long accum a)
short fract __fractqiqq (signed char a)
fract __fractqihq (signed char a)
long fract __fractqisq (signed char a)
long long fract __fractqidq (signed char a)
short accum __fractqiha (signed char a)
accum __fractqisa (signed char a)
long accum __fractqiuda (signed char a)
long long accum __fractqiuta (signed char a)
unsigned short fract __fractqiussq (signed char a)
unsigned fract __fractqiuhq (signed char a)
unsigned long long fract __fractqiuidq (signed char a)
unsigned short accum __fractqiuhla (signed char a)
unsigned accum __fractqiusa (signed char a)
unsigned long accum __fractqiuuda (signed char a)
unsigned long long accum __fractqiuda (signed char a)
short fract __fracthiqq (short a)
fract __fracthihq (short a)
long fract __fracthisq (short a)
long long fract __fracthidq (short a)
short accum __fracthiha (short a)
accum __fracthisa (short a)
long accum __fracthida (short a)
long long accum __fracthita (short a)
unsigned short fract __fracthiuqq (short a)
unsigned fract __fracthiuhq (short a)
unsigned long long fract __fracthiuidq (short a)
unsigned short accum __fracthiuha (short a)
unsigned accum __fracthiusa (short a)
unsigned long accum __fracthida (short a)
unsigned long long accum __fracthita (short a)
short fract __fractsiqq (int a)
fract __fractsihq (int a)
l
short accum __fractsiha (int a) [Runtime Function]
accum __fractsisa (int a) [Runtime Function]
long accum __fractsida (int a) [Runtime Function]
long long accum __fractsita (int a) [Runtime Function]
unsigned short fract __fractsiuqq (int a) [Runtime Function]
unsigned fract __fractsiuhq (int a) [Runtime Function]
unsigned long fract __fractsiusq (int a) [Runtime Function]
unsigned long long fract __fractsitdq (int a) [Runtime Function]
unsigned short accum __fractsiuha (int a) [Runtime Function]
unsigned long accum __fractsiusa (int a) [Runtime Function]
unsigned long long accum __fractsiuta (int a) [Runtime Function]
short fract __fractdiqq (long a) [Runtime Function]
fract __fractdihq (long a) [Runtime Function]
long fract __fractdisq (long a) [Runtime Function]
long long fract __fractdisq (long a) [Runtime Function]
short accum __fractdiha (long a) [Runtime Function]
accum __fractdisa (long a) [Runtime Function]
long accum __fractdida (long a) [Runtime Function]
long long accum __fractdita (long a) [Runtime Function]
unsigned short fract __fractdiuqq (long a) [Runtime Function]
unsigned fract __fractdiuhq (long a) [Runtime Function]
unsigned long fract __fractdiusq (long a) [Runtime Function]
unsigned long long fract __fractdiudq (long a) [Runtime Function]
unsigned short accum __fractdiuha (long a) [Runtime Function]
unsigned long accum __fractdiusa (long a) [Runtime Function]
unsigned long long accum __fractdiuta (long a) [Runtime Function]
short fract __fracttiqq (long long a) [Runtime Function]
fact __fracttihq (long long a) [Runtime Function]
long fract __fracttisq (long long a) [Runtime Function]
long long fract __fracttisq (long long a) [Runtime Function]
short accum __fracttih (long long a) [Runtime Function]
accum __fracttisa (long long a) [Runtime Function]
long accum __fracttida (long long a) [Runtime Function]
long long accum __fracttita (long long a) [Runtime Function]
unsigned short fract __fracttiuqq (long long a) [Runtime Function]
unsigned fract __fracttiuhq (long long a) [Runtime Function]
unsigned long fract __fracttiusq (long long a) [Runtime Function]
unsigned long long fract __fracttiudq (long long a) [Runtime Function]
unsigned short accum __fracttiuha (long long a) [Runtime Function]
unsigned accm __fracttiusa (long long a) [Runtime Function]
unsigned long accum __fracttiuda (long long a) [Runtime Function]
unsigned long long accum __fracttiuta (long long a) [Runtime Function]
short fract __fractsfqq (float a) [Runtime Function]
fract __fractsfhq (float a) [Runtime Function]
long fract __fractsfssq (float a) [Runtime Function]
These functions convert from fractional and signed non-fractionals to fractionals and signed non-fractionals, without saturation.
Chapter 4: The GCC low-level runtime library

Runtime Function

unsigned long long accum __satfractdqua(long long fract a)

Runtime Function

short fract __satfractshaqq(short accum a)

Runtime Function

fract __satfracthahq(short accum a)

Runtime Function

long fract __satfracthasq(short accum a)

Runtime Function

long long fract __satfracthadq(short accum a)

Runtime Function

accum __satfracthasa2(short accum a)

Runtime Function

long accum __satfracthada2(short accum a)

Runtime Function

long long accum __satfracthata2(short accum a)

Runtime Function

unsigned short fract __satfracthauqq(short accum a)

Runtime Function

unsigned fract __satfracthauhq(short accum a)

Runtime Function

unsigned long fract __satfracthausq(short accum a)

Runtime Function

unsigned long long fract __satfracthaudq(short accum a)

Runtime Function

unsigned short accum __satfracthauha(short accum a)

Runtime Function

unsigned accum __satfracthausa(short accum a)

Runtime Function

unsigned long accum __satfracthauda(short accum a)

Runtime Function

unsigned long long accum __satfracthauta(short accum a)

Runtime Function

short fract __satfractsaqq(accum a)

Runtime Function

fract __satfractsahq(accum a)

Runtime Function

long fract __satfractsasq(accum a)

Runtime Function

long long fract __satfractsadq(accum a)

Runtime Function

short accum __satfractsaha2(accum a)

Runtime Function

long accum __satfractsada2(accum a)

Runtime Function

long long accum __satfractsata2(accum a)

Runtime Function

unsigned short fract __satfractsauqq(accum a)

Runtime Function

unsigned fract __satfractsauhq(accum a)

Runtime Function

unsigned long fract __satfractsausq(accum a)

Runtime Function

unsigned long long fract __satfractsaudq(accum a)

Runtime Function

unsigned short accum __satfractsauha(accum a)

Runtime Function

unsigned accum __satfractsausa(accum a)

Runtime Function

unsigned long accum __satfractsauda(accum a)

Runtime Function

unsigned long long accum __satfractsauta(accum a)

Runtime Function

short fract __satfractdaqq(long accum a)

Runtime Function

fract __satfractdahq(long accum a)

Runtime Function

long fract __satfractdasq(long accum a)

Runtime Function

long long fract __satfractdadq(long accum a)

Runtime Function

short accum __satfractdaha2(long accum a)

Runtime Function

accum __satfractdasa2(long accum a)

Runtime Function

long long accum __satfractdata2(long accum a)
unsigned accum __satfractdausa (long accum a) [Runtime Function]
unsigned long accum __satfractdauda (long accum a) [Runtime Function]
unsigned long long accum __satfractdauta (long accum a) [Runtime Function]
short fract __satfracttaqq (long long accum a) [Runtime Function]
fract __satfracttahq (long long accum a) [Runtime Function]
long fract __satfracttasq (long long accum a) [Runtime Function]
long long fract __satfracttdaq (long long accum a) [Runtime Function]
short accum __satfracttaha2 (long long accum a) [Runtime Function]
accum __satfracttasaa2 (long long accum a) [Runtime Function]
long accum __satfracttada2 (long long accum a) [Runtime Function]
unsigned short fract __satfracttauqq (long long accum a) [Runtime Function]
unsigned fract __satfracttauhq (long long accum a) [Runtime Function]
unsigned long fract __satfracttausq (long long accum a) [Runtime Function]
unsigned long long fract __satfracttaudq (long long accum a) [Runtime Function]
unsigned short accum __satfracttahuha (long long accum a) [Runtime Function]
unsigned accum __satfracttausa (long long accum a) [Runtime Function]
unsigned long accum __satfracttauda (long long accum a) [Runtime Function]
unsigned long long accum __satfracttauta (long long accum a) [Runtime Function]
short fract __satfractuqqqq (unsigned short fract a) [Runtime Function]
fract __satfractuqqhq (unsigned short fract a) [Runtime Function]
long fract __satfractuqqsq (unsigned short fract a) [Runtime Function]
long long fract __satfractuqqdq (unsigned short fract a) [Runtime Function]
short accum __satfractuqqha (unsigned short fract a) [Runtime Function]
accum __satfractuqqsa (unsigned short fract a) [Runtime Function]
long accum __satfractuqqda (unsigned short fract a) [Runtime Function]
long long accum __satfractuqqta (unsigned short fract a) [Runtime Function]
unsigned fract __satfractuqquhq2 (unsigned short fract a) [Runtime Function]
unsigned long fract __satfractuqqusq2 (unsigned short fract a) [Runtime Function]
unsigned long long fract __satfractuqqudq2 (unsigned short fract a) [Runtime Function]
unsigned short accum __satfractuqqhua (unsigned short fract a) [Runtime Function]
unsigned accum __satfractuqqusa (unsigned short fract a) [Runtime Function]
unsigned long accum __satfractuqquda (unsigned short fract a) [Runtime Function]
unsigned long long accum __satfractuqquta (unsigned short fract a) [Runtime Function]
short fract __satfractuhqqq (unsigned fract a) [Runtime Function]
fract __satfractuhhq (unsigned fract a) [Runtime Function]
long fract __satfractuhhq (unsigned fract a) [Runtime Function]
long long fract __satfractuhhq (unsigned fract a) [Runtime Function]
<table>
<thead>
<tr>
<th>Function Name</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>short accum __satfractuhqha</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>accum __satfractuhqsa</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long accum __satfractuhqda</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long long accum __satfractuhqta</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned short fract __satfractuhquqq2</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned long fract __satfractuhqusq2</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned short accum __satfractuhquha</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned accum __satfractuhqasa</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned long accum __satfractuhquda</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned long long accum __satfractuhquta</td>
<td>(unsigned fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>short fract __satfractusqqq</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>fract __satfractusqqq</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long fract __satfractusqsq</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long long fract __satfractusqqq</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>short accum __satfractusqa</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long accum __satfractusqda</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long long accum __satfractusqta</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned short fract __satfractusquqq2</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned long long fract __satfractusqudq2</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>short fract __satfractusquha</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned accum __satfractusqusa</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned long accum __satfractusquda</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>unsigned long long accum __satfractusquta</td>
<td>(unsigned long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>short fract __satfractudqqq</td>
<td>(unsigned long long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>fract __satfractudqqq</td>
<td>(unsigned long long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long fract __satfractudqsq</td>
<td>(unsigned long long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long long fract __satfractudqqq</td>
<td>(unsigned long long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>short accum __satfractudqha</td>
<td>(unsigned long long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>accum __satfractudqsa</td>
<td>(unsigned long long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long accum __satfractudqda</td>
<td>(unsigned long long fract a)</td>
<td>Runtime Function</td>
</tr>
<tr>
<td>long long accum __satfractudqta</td>
<td>(unsigned long long fract a)</td>
<td>Runtime Function</td>
</tr>
</tbody>
</table>
unsigned short fract __satfractudquqq2 (unsigned long long fract a) [Runtime Function]
unsigned fract __satfractudquhq2 (unsigned long long fract a) [Runtime Function]
unsigned long fract __satfractudqusq2 (unsigned long long fract a) [Runtime Function]
unsigned short accum __satfractudquha (unsigned long long fract a) [Runtime Function]
unsigned accum __satfractudquasa (unsigned long long fract a) [Runtime Function]
unsigned long accum __satfractudquda (unsigned long long fract a) [Runtime Function]
unsigned long long accum __satfractudquhata (unsigned long long fract a) [Runtime Function]
short fract __satfractuhaha (unsigned short accum a) [Runtime Function]
fract __satfractuhahq (unsigned short accum a) [Runtime Function]
long fract __satfractuhasaq (unsigned short accum a) [Runtime Function]
long long fract __satfractuhadq (unsigned short accum a) [Runtime Function]
short accum __satfractuhaha (unsigned short accum a) [Runtime Function]
accum __satfractuhahsq (unsigned short accum a) [Runtime Function]
long accum __satfractuhadq (unsigned short accum a) [Runtime Function]
long long accum __satfractuhadq (unsigned short accum a) [Runtime Function]
unsigned short fract __satfractuhauqq (unsigned short accum a) [Runtime Function]
unsigned fract __satfractuhauhq (unsigned short accum a) [Runtime Function]
unsigned long fract __satfractuhasq (unsigned short accum a) [Runtime Function]
unsigned long long fract __satfractuhadq (unsigned short accum a) [Runtime Function]
unsigned accum __satfractuhasa2 (unsigned short accum a) [Runtime Function]
unsigned long accum __satfractuhada2 (unsigned short accum a) [Runtime Function]
unsigned long long accum __satfractuhada2 (unsigned short accum a) [Runtime Function]
short fract __satfractusaqq (unsigned accum a) [Runtime Function]
fract __satfractusaq (unsigned accum a) [Runtime Function]
long fract __satfractusaq (unsigned accum a) [Runtime Function]
long long fract __satfractusaq (unsigned accum a) [Runtime Function]
short accum __satfractusaq (unsigned accum a) [Runtime Function]
accum __satfractusaq (unsigned accum a) [Runtime Function]
long accum __satfractusaq (unsigned accum a) [Runtime Function]
long long accum __satfractusaq (unsigned accum a) [Runtime Function]
unsigned short fract __satfractusauqq (unsigned accum a) [Runtime Function]
unsigned fract __satfractusauhq (unsigned accum a) [Runtime Function]
unsigned long fract __satfractusausq (unsigned accum a) [Runtime Function]
unsigned long long fract __satfractusaudq (unsigned accum a) [Runtime Function]
unsigned short accum __satfractusauha2 (unsigned accum a) [Runtime Function]
unsigned long accum __satfractusauha2 (unsigned accum a) [Runtime Function]
unsigned long long accum __satfractusauha2 (unsigned accum a) [Runtime Function]
short fract __satfractudaqq (unsigned long accum a) [Runtime Function]
fract __satfractudahq (unsigned long accum a) [Runtime Function]
long fract __satfractudaq (unsigned long accum a) [Runtime Function]
long long fract __satfractudadq (unsigned long accum a) [Runtime Function]
short accum __satfractudaha (unsigned long accum a) [Runtime Function]
accum __satfractudaha (unsigned long accum a) [Runtime Function]
long accum __satfractudada (unsigned long accum a) [Runtime Function]
long long accum __satfractudata (unsigned long accum a) [Runtime Function]
unsigned short fract __satfractudaqq (unsigned long accum a) [Runtime Function]
unsigned fract __satfractudaqq (unsigned long accum a) [Runtime Function]
unsigned long fract __satfractudaqq (unsigned long accum a) [Runtime Function]
unsigned long long fract __satfractudaqq (unsigned long accum a) [Runtime Function]
short fract __satfractudaqq (unsigned long accum a) [Runtime Function]
fract __satfractudaqq (unsigned long accum a) [Runtime Function]
long fract __satfractudaqq (unsigned long accum a) [Runtime Function]
long long fract __satfractudaqq (unsigned long accum a) [Runtime Function]
short accum __satfractudaqq (unsigned long accum a) [Runtime Function]
accum __satfractudaqq (unsigned long accum a) [Runtime Function]
long accum __satfractudaqq (unsigned long accum a) [Runtime Function]
long long accum __satfractudaqq (unsigned long accum a) [Runtime Function]
unsigned short fract __satfractudaqq (unsigned long accum a) [Runtime Function]
unsigned fract __satfractutauhq (unsigned long long
  accum a)
unsigned long fract __satfractutausq (unsigned long
  long accum a)
unsigned long long fract __satfractutaudq (unsigned
  long long accum a)
unsigned short accum __satfractutauha2 (unsigned long
  long accum a)
unsigned accum __satfractutausa2 (unsigned long long
  accum a)
unsigned long accum __satfractutauda2 (unsigned long
  long accum a)
short fract __satfractqiqq (signed char a)
fract __satfractqihq (signed char a)
long fract __satfractqisq (signed char a)
long long fract __satfractqidq (signed char a)
short accum __satfractqiha (signed char a)
accum __satfractqisa (signed char a)
long accum __satfractqiida (signed char a)
long long accum __satfractqiita (signed char a)
unsigned short fract __satfractqiuqq (signed char a)
unsigned fract __satfractqiuhq (signed char a)
unsigned long fract __satfractqiusq (signed char a)
unsigned long long fract __satfractqiudq (signed char
  a)
unsigned short accum __satfractqiua (signed char a)
unsigned accum __satfractqiua (signed char a)
unsigned long accum __satfractqiuda (signed char a)
unsigned long long accum __satfractqiuta (signed char
  a)
short fract __satfracthiqq (short a)
fract __satfracthihq (short a)
long fract __satfracthisq (short a)
long long fract __satfracthidq (short a)
short accum __satfracthiha (short a)
accum __satfracthisa (short a)
long accum __satfracthida (short a)
long long accum __satfracthita (short a)
unsigned short fract __satfracthiuqq (short a)
unsigned fract __satfracthiuhq (short a)
unsigned long fract __satfracthiusq (short a)
unsigned long long fract __satfracthiuq (short a)
unsigned short accum __satfracthiua (short a)
unsigned accum __satfracthiusa (short a)
unsigned long accum __satfracthiuda (short a)
unsigned long long accum __satfracthiuta (short a)
short fract __satfractsiqq (int a)
fract __satfractsihq (int a)
long fract __satfractsiq (int a)
long long fract __satfractsidq (int a)
short accum __satfractsisa (int a)
long accum __satfractsida (int a)
long long accum __satfractsitaua (int a)
unsigned short fract __satfractsiusq (int a)
unsigned fract __satfractsiuhq (int a)
unsigned long fract __satfractsiusq (int a)
unsigned long long fract __satfractsiuudq (int a)
unsigned short accum __satfractsiuha (int a)
unsigned accum __satfractsiusa (int a)
unsigned long accum __satfractsiuda (int a)
unsigned long long accum __satfractsiuta (int a)
short fract __satfractdiquq (long a)
short long fract __satfractdiuqq (long a)
fract __satfractdihq (long a)
long fract __satfractdisq (long a)
long long fract __satfractdisq (long a)
short accum __satfractdiha (long a)
accum __satfractdisa (long a)
long accum __satfractdisa (long a)
long long accum __satfractdita (long a)
unsigned short fract __satfractdiuqq (long a)
unsigned fract __satfractdiuhq (long a)
unsigned long fract __satfractdiuqq (long a)
unsigned long long fract __satfractdiuudq (long a)
unsigned short accum __satfractdiua (long a)
unsigned accum __satfractdiusa (long a)
unsigned long accum __satfractdiuda (long a)
unsigned long long accum __satfractdiuta (long a)
short fract __satfracttiq (long long a)
fract __satfracttihq (long long a)
long fract __satfracttisq (long long a)
long long fract __satfracttisq (long long a)
short accum __satfracttiha (long long a)
accum __satfracttisa (long long a)
long accum __satfracttida (long long a)
long long accum __satfracttida (long long a)
unsigned short fract __satfracttiuq (long long a)
unsigned long fract __satfracttiuq (long long a)
unsigned long long fract __satfracttiuudq (long long a)
unsigned short accum __satfractti uhq (long long a)
unsigned accum __satfracttiusa (long long a)
unsigned long accum __satfracttiuda (long long a)
unsigned long long accum __satfracttiuta (long long a)
The functions convert from fractional and signed non-fractionals to fractionals, with saturation.
unsigned int __fractunssqsi (long fract a) [Runtime Function]
unsigned long __fractunssqdi (long fract a) [Runtime Function]
unsigned long long __fractunssqti (long fract a) [Runtime Function]
unsigned char __fractunssdqqi (long fract a) [Runtime Function]
unsigned short __fractunssdqhi (long fract a) [Runtime Function]
unsigned int __fractunssdqsi (long fract a) [Runtime Function]
unsigned long __fractunssdqdi (long fract a) [Runtime Function]
unsigned long long __fractunssdqti (long fract a) [Runtime Function]
unsigned char __fractunshaqi (short accum a) [Runtime Function]
unsigned short __fractunshahi (short accum a) [Runtime Function]
unsigned int __fractunshasi (short accum a) [Runtime Function]
unsigned long __fractunshadi (short accum a) [Runtime Function]
unsigned long long __fractunshati (short accum a) [Runtime Function]
unsigned char __fractunshaqi (accum a) [Runtime Function]
unsigned short __fractunshahi (accum a) [Runtime Function]
unsigned int __fractunshasi (accum a) [Runtime Function]
unsigned long __fractunssadai (accum a) [Runtime Function]
unsigned long long __fractunssatari (accum a) [Runtime Function]
unsigned char __fractunshadi (long accum a) [Runtime Function]
unsigned short __fractunshahi (long accum a) [Runtime Function]
unsigned int __fractunshasi (long accum a) [Runtime Function]
unsigned long __fractunshadi (long accum a) [Runtime Function]
unsigned long long __fractunshati (long accum a) [Runtime Function]
unsigned char __fractunshadi (long long accum a) [Runtime Function]
unsigned short __fractunshahi (long long accum a) [Runtime Function]
unsigned int __fractunshasi (long long accum a) [Runtime Function]
unsigned long __fractunshadi (long long accum a) [Runtime Function]
unsigned long long __fractunshati (long long accum a) [Runtime Function]
unsigned char __fractunssqhi (unsigned short fract a) [Runtime Function]
unsigned short __fractunssqhi (unsigned short fract a) [Runtime Function]
unsigned int __fractunssqsi (unsigned short fract a) [Runtime Function]
unsigned long __fractunssqdi (unsigned short fract a) [Runtime Function]
unsigned long long __fractunssqti (unsigned short fract a) [Runtime Function]
unsigned char __fractunssqhi (unsigned long fract a) [Runtime Function]
unsigned short __fractunssqhi (unsigned long fract a) [Runtime Function]
unsigned int __fractunssqsi (unsigned long fract a) [Runtime Function]
unsigned long __fractunssqdi (unsigned long fract a) [Runtime Function]
unsigned long long __fractunssqti (unsigned long fract a) [Runtime Function]
unsigned char __fractunssdqqi (unsigned long long fract a) [Runtime Function]
unsigned short __fractunssdqqi (unsigned long long fract a) [Runtime Function]
unsigned int __fractunssdqsi (unsigned long long fract a) [Runtime Function]
unsigned long __fractunssdqdi (unsigned long long fract a) [Runtime Function]
unsigned long long __fractunssdqti (unsigned long long fract a) [Runtime Function]
unsigned accum __fractunsqiusa (unsigned char a)
unsigned long accum __fractunsqiuda (unsigned char a)
unsigned long long accum __fractunsqiuta (unsigned char a)

short fract __fractunshiqq (unsigned short a)
fract __fractunshihq (unsigned short a)
long fract __fractunshisq (unsigned short a)
long long fract __fractunshidq (unsigned short a)
short accum __fractunshiha (unsigned short a)
accum __fractunshisa (unsigned short a)
long accum __fractunshida (unsigned short a)
long long accum __fractunshita (unsigned short a)
unsigned short fract __fractunshiuqq (unsigned short a)
unsigned fract __fractunshiuhq (unsigned short a)
unsigned long fract __fractunshiusq (unsigned short a)
unsigned long long fract __fractunshidq (unsigned short a)
unsigned short accum __fractunshiha (unsigned short a)
unsigned accum __fractunshisa (unsigned short a)
unsigned long accum __fractunshida (unsigned short a)
unsigned long long accum __fractunshita (unsigned short a)

short fract __fractunssiqq (unsigned int a)
fract __fractunssihq (unsigned int a)
long fract __fractunssisq (unsigned int a)
long long fract __fractunssidq (unsigned int a)
short accum __fractunssihq (unsigned int a)
accum __fractunssisa (unsigned int a)
long accum __fractunssida (unsigned int a)
long long accum __fractunssita (unsigned int a)
unsigned short fract __fractunssiuqq (unsigned int a)
unsigned fract __fractunssiuhq (unsigned int a)
unsigned long fract __fractunssiusq (unsigned int a)
unsigned long long fract __fractunssidq (unsigned int a)
unsigned short accum __fractunssiha (unsigned int a)
unsigned accum __fractunssisa (unsigned int a)
unsigned long accum __fractunssida (unsigned int a)
unsigned long long accum __fractunssita (unsigned int a)

short fract __fractunsdiqq (unsigned long a)
fract __fractunsdihq (unsigned long a)
long fract __fractunsdisq (unsigned long a)
long long fract __fractunsdidq (unsigned long a)
short accum __fractunsdiha (unsigned long a)
accum __fractunsdisa (unsigned long a)
long accum __fractunsdida (unsigned long a)
These functions convert from fractionals to unsigned non-fractionals; and from unsigned non-fractionals to fractionals, without saturation.
unsigned long fract __satfractunsqiusq (unsigned char a) [Runtime Function]
unsigned long long fract __satfractunsqiudq (unsigned char a) [Runtime Function]
unsigned short accum __satfractunsqiua (unsigned char a) [Runtime Function]
unsigned accum __satfractunsqiusa (unsigned char a) [Runtime Function]
unsigned long accum __satfractunsqiuda (unsigned char a) [Runtime Function]
unsigned long long accum __satfractunsqiuta (unsigned char a) [Runtime Function]
short fract __satfractunshiqq (unsigned short a) [Runtime Function]
fract __satfractunshihq (unsigned short a) [Runtime Function]
long fract __satfractunshisq (unsigned short a) [Runtime Function]
long long fract __satfractunshidq (unsigned short a) [Runtime Function]
short accum __satfractunshihua (unsigned short a) [Runtime Function]
accum __satfractunshisua (unsigned short a) [Runtime Function]
long accum __satfractunshida (unsigned short a) [Runtime Function]
long long accum __satfractunshidua (unsigned short a) [Runtime Function]
unsigned short fract __satfractunshiuqq (unsigned short a) [Runtime Function]
unsigned long fract __satfractunshiuqh (unsigned short a) [Runtime Function]
unsigned long long fract __satfractunshiusq (unsigned short a) [Runtime Function]
unsigned short accum __satfractunshihu (unsigned short a) [Runtime Function]
unsigned accum __satfractunshis (unsigned short a) [Runtime Function]
unsigned long accum __satfractunshida (unsigned short a) [Runtime Function]
unsigned long long accum __satfractunshida (unsigned short a) [Runtime Function]
short fract __satfractunssiqq (unsigned int a) [Runtime Function]
fract __satfractunssihq (unsigned int a) [Runtime Function]
long fract __satfractunssisq (unsigned int a) [Runtime Function]
long long fract __satfractunssidq (unsigned int a) [Runtime Function]
short accum __satfractunssihua (unsigned int a) [Runtime Function]
accum __satfractunssisua (unsigned int a) [Runtime Function]
long accum __satfractunssida (unsigned int a) [Runtime Function]
long long accum __satfractunssidua (unsigned int a) [Runtime Function]
unsigned short fract __satfractunssiuqq (unsigned int a) [Runtime Function]
unsigned long fract __satfractunssiuh (unsigned int a) [Runtime Function]
unsigned long long fract __satfractunssiusq (unsigned int a) [Runtime Function]
unsigned long long fract __satfractunssiudq (unsigned int a) [Runtime Function]
unsigned short accum __satfractunssiua (unsigned int a) [Runtime Function]
unsigned accum __satfractunssiusa (unsigned int a) [Runtime Function]
unsigned long accum __satfractunssiuda (unsigned int a) [Runtime Function]
unsigned long long accum __satfractunssiuta (unsigned int a) [Runtime Function]
short fract __satfractunssiudq (unsigned long a) [Runtime Function]
fract __satfractunssihuq (unsigned long a) [Runtime Function]
long fract __satfractunssidq (unsigned long a) [Runtime Function]
long long fract __satfractunssidq (unsigned long a) [Runtime Function]
short accum __satfractunssihuq (unsigned long a) [Runtime Function]
accum __satfractunssisuq (unsigned long a) [Runtime Function]
long accum __satfractunssiuha (unsigned long a) [Runtime Function]
long long accum __satfractunssiusa (unsigned long a) [Runtime Function]
unsigned short fract __satfractunssiudq (unsigned long a) [Runtime Function]
unsigned fract __satfractunssihuq (unsigned long long a) [Runtime Function]
unsigned long fract __satfractunssiusq (unsigned long a) [Runtime Function]
unsigned long fract __satfractunssiudq (unsigned long a) [Runtime Function]
unsigned short accum __satfractunssiua (unsigned long a) [Runtime Function]
unsigned accum __satfractunssiusa (unsigned long a) [Runtime Function]
unsigned long accum __satfractunssiuda (unsigned long a) [Runtime Function]
unsigned long long accum __satfractunssiuta (unsigned long a) [Runtime Function]
short fract __satfractunssiudq (unsigned long long a) [Runtime Function]
fract __satfractunstihuq (unsigned long long a) [Runtime Function]
long fract __satfractunstisq (unsigned long long a) [Runtime Function]
long long fract __satfractunstidq (unsigned long long a) [Runtime Function]
short accum __satfractunstihu (unsigned long long a) [Runtime Function]
accum __satfractunstisu (unsigned long long a) [Runtime Function]
long accum __satfractunstida (unsigned long long a) [Runtime Function]
long long accum __satfractunstidq (unsigned long long a) [Runtime Function]
unsigned short fract __satfractunstiuq (unsigned long long a) [Runtime Function]
unsigned fract __satfractunstihuq (unsigned long long a) [Runtime Function]
unsigned long fract __satfractunstiusq (unsigned long long a) [Runtime Function]
unsigned long long fract __satfractunstiudq (unsigned long long a) [Runtime Function]
4.5 Language-independent routines for exception handling

document me!

_Unwind_DeleteException
_Unwind_Find_FDE
_Unwind_ForcedUnwind
_Unwind_GetGR
_Unwind_GetIP
_Unwind_GetLanguageSpecificData
_Unwind_GetRegionStart
_Unwind_GetTextRelBase
_Unwind_GetDataRelBase
_Unwind_RaiseException
_Unwind_Resume
_Unwind_SetGR
_Unwind_SetIP
_Unwind_FindEnclosingFunction
_Unwind_SjLj_Register
_Unwind_SjLj_Unregister
_Unwind_SjLj_RaiseException
_Unwind_SjLj_ForcedUnwind
_Unwind_SjLj_Resume
__deregister_frame
__deregister_frame_info
__deregister_frame_info_bases
__register_frame
__register_frame_info
__register_frame_info_bases
__register_frame_info_table
__register_frame_info_table_bases
__register_frame_table

4.6 Miscellaneous runtime library routines

4.6.1 Cache control functions

void __clear_cache (char *beg, char *end) [Runtime Function]
This function clears the instruction cache between beg and end.

4.6.2 Split stack functions and variables

void * __splitstack_find (void *segment_arg, void *sp, size_t len, void **next_segment, void **next_sp, void **initial_sp) [Runtime Function]
When using ‘-fsplit-stack’, this call may be used to iterate over the stack segments.
It may be called like this:
void *next_segment = NULL;
void *next_sp = NULL;
void *initial_sp = NULL;
void *stack;
size_t stack_size;
while ((stack = __splitstack_find (next_segment, next_sp,
&stack_size, &next_segment,
&next_sp, &initial_sp))
!= NULL)
{
    /* Stack segment starts at stack and is
     stack_size bytes long. */
}

There is no way to iterate over the stack segments of a different thread. However,
what is permitted is for one thread to call this with the segment_arg and sp arguments NULL, to pass next_segment, next_sp, and initial_sp to a different thread, and
then to suspend one way or another. A different thread may run the subsequent
__splitstack_find iterations. Of course, this will only work if the first thread is
suspended while the second thread is calling __splitstack_find. If not, the second
thread could be looking at the stack while it is changing, and anything could happen.

__morestack_segments  [Variable]
__morestack_current_segment [Variable]
__morestack_initial_sp  [Variable]

Internal variables used by the ‘-fsplit-stack’ implementation.
Chapter 5: Language Front Ends in GCC

The interface to front ends for languages in GCC, and in particular the tree structure (see Chapter 11 [GENERIC], page 177), was initially designed for C, and many aspects of it are still somewhat biased towards C and C-like languages. It is, however, reasonably well suited to other procedural languages, and front ends for many such languages have been written for GCC.

Writing a compiler as a front end for GCC, rather than compiling directly to assembler or generating C code which is then compiled by GCC, has several advantages:

- GCC front ends benefit from the support for many different target machines already present in GCC.
- GCC front ends benefit from all the optimizations in GCC. Some of these, such as alias analysis, may work better when GCC is compiling directly from source code then when it is compiling from generated C code.
- Better debugging information is generated when compiling directly from source code than when going via intermediate generated C code.

Because of the advantages of writing a compiler as a GCC front end, GCC front ends have also been created for languages very different from those for which GCC was designed, such as the declarative logic/functional language Mercury. For these reasons, it may also be useful to implement compilers created for specialized purposes (for example, as part of a research project) as GCC front ends.
6 Source Tree Structure and Build System

This chapter describes the structure of the GCC source tree, and how GCC is built. The user documentation for building and installing GCC is in a separate manual (http://gcc.gnu.org/install/), with which it is presumed that you are familiar.

6.1 Configure Terms and History

The configure and build process has a long and colorful history, and can be confusing to anyone who doesn’t know why things are the way they are. While there are other documents which describe the configuration process in detail, here are a few things that everyone working on GCC should know.

There are three system names that the build knows about: the machine you are building on (build), the machine that you are building for (host), and the machine that GCC will produce code for (target). When you configure GCC, you specify these with ‘--build=’, ‘--host=’, and ‘--target=’.

Specifying the host without specifying the build should be avoided, as configure may (and once did) assume that the host you specify is also the build, which may not be true.

If build, host, and target are all the same, this is called a native. If build and host are the same but target is different, this is called a cross. If build, host, and target are all different this is called a canadian (for obscure reasons dealing with Canada’s political party and the background of the person working on the build at that time). If host and target are the same, but build is different, you are using a cross-compiler to build a native for a different system. Some people call this a host-x-host, crossed native, or cross-built native. If build and target are the same, but host is different, you are using a cross compiler to build a cross compiler that produces code for the machine you’re building on. This is rare, so there is no common way of describing it. There is a proposal to call this a crossback.

If build and host are the same, the GCC you are building will also be used to build the target libraries (like libstdc++). If build and host are different, you must have already built and installed a cross compiler that will be used to build the target libraries (if you configured with ‘--target=foo-bar’, this compiler will be called foo-bar-gcc).

In the case of target libraries, the machine you’re building for is the machine you specified with ‘--target’. So, build is the machine you’re building on (no change there), host is the machine you’re building for (the target libraries are built for the target, so host is the target you specified), and target doesn’t apply (because you’re not building a compiler, you’re building libraries). The configure/make process will adjust these variables as needed. It also sets $with_cross_host to the original ‘--host’ value in case you need it.

The libiberty support library is built up to three times: once for the host, once for the target (even if they are the same), and once for the build if build and host are different. This allows it to be used by all programs which are generated in the course of the build process.

6.2 Top Level Source Directory

The top level source directory in a GCC distribution contains several files and directories that are shared with other software distributions such as that of GNU Binutils. It also contains several subdirectories that contain parts of GCC and its runtime libraries:
'boehm-gc'
   The Boehm conservative garbage collector, used as part of the Java runtime library.

'config'
   Autoconf macros and Makefile fragments used throughout the tree.

'contrib'
   Contributed scripts that may be found useful in conjunction with GCC. One of these, 'contrib/texi2pod.pl', is used to generate man pages from Texinfo manuals as part of the GCC build process.

'fixincludes'
   The support for fixing system headers to work with GCC. See 'fixincludes/README' for more information. The headers fixed by this mechanism are installed in 'libsubdir/include-fixed'. Along with those headers, 'README-fixinc' is also installed, as 'libsubdir/include-fixed/README'.

'gcc'
   The main sources of GCC itself (except for runtime libraries), including optimizers, support for different target architectures, language front ends, and testsuites. See Section 6.3 [The 'gcc' Subdirectory], page 63, for details.

'gnattools'
   Support tools for GNAT.

'include'
   Headers for the libiberty library.

'intl'
   GNU libintl, from GNU gettext, for systems which do not include it in libc.

'libada'
   The Ada runtime library.

'libcpp'
   The C preprocessor library.

'libdecnumber'
   The Decimal Float support library.

'libffi'
   The libffi library, used as part of the Java runtime library.

'libgcc'
   The GCC runtime library.

'libfortran'
   The Fortran runtime library.

'libgo'
   The Go runtime library. The bulk of this library is mirrored from the master Go repository.

'libgomp'
   The GNU OpenMP runtime library.

'libiberty'
   The libiberty library, used for portability and for some generally useful data structures and algorithms. See Section “Introduction” in GNU libiberty, for more information about this library.

'libjava'
   The Java runtime library.

'libmudflap'
   The libmudflap library, used for instrumenting pointer and array dereferencing operations.
‘libobjc’ The Objective-C and Objective-C++ runtime library.

‘libssp’ The Stack protector runtime library.

‘libstdc++-v3’
The C++ runtime library.

‘lto-plugin’
Plugin used by gold if link-time optimizations are enabled.

‘maintainer-scripts’
Scripts used by the gccadmin account on gcc.gnu.org.

‘zlib’ The zlib compression library, used by the Java front end, as part of the Java runtime library, and for compressing and uncompressing GCC’s intermediate language in LTO object files.

The build system in the top level directory, including how recursion into subdirectories works and how building runtime libraries for multilibs is handled, is documented in a separate manual, included with GNU Binutils. See Section “GNU configure and build system” in The GNU configure and build system, for details.

6.3 The ‘gcc’ Subdirectory

The ‘gcc’ directory contains many files that are part of the C sources of GCC, other files used as part of the configuration and build process, and subdirectories including documentation and a test suite. The files that are sources of GCC are documented in a separate chapter. See Chapter 9 [Passes and Files of the Compiler], page 111.

6.3.1 Subdirectories of ‘gcc’

The ‘gcc’ directory contains the following subdirectories:

‘language’
Subdirectories for various languages. Directories containing a file ‘config-lang.in’ are language subdirectories. The contents of the subdirectories ‘cp’ (for C++), ‘lto’ (for LTO), ‘objc’ (for Objective-C) and ‘objc++’ (for Objective-C++) are documented in this manual (see Chapter 9 [Passes and Files of the Compiler], page 111); those for other languages are not. See Section 6.3.8 [Anatomy of a Language Front End], page 71, for details of the files in these directories.

‘config’ Configuration files for supported architectures and operating systems. See Section 6.3.9 [Anatomy of a Target Back End], page 75, for details of the files in this directory.

‘doc’ Texinfo documentation for GCC, together with automatically generated man pages and support for converting the installation manual to HTML. See Section 6.3.7 [Documentation], page 69.

‘ginclude’ System headers installed by GCC, mainly those required by the C standard of freestanding implementations. See Section 6.3.6 [Headers Installed by GCC], page 68, for details of when these and other headers are installed.
Message catalogs with translations of messages produced by GCC into various languages, \texttt{language.po}. This directory also contains \texttt{gcc.pot}, the template for these message catalogues, \texttt{exgettext}, a wrapper around \texttt{gettext} to extract the messages from the GCC sources and create \texttt{gcc.pot}, which is run by \texttt{make gcc.pot}, and \texttt{EXCLUDES}, a list of files from which messages should not be extracted.

The GCC testsuites (except for those for runtime libraries). See Chapter 7 [Testsuites], page 77.

6.3.2 Configuration in the `gcc` Directory

The `gcc` directory is configured with an Autoconf-generated script `configure`. The `configure` script is generated from `configure.ac` and `aclocal.m4`. From the files `configure.ac` and `acconfig.h`, Autoheader generates the file `config.in`. The file `cstamp-h.in` is used as a timestamp.

6.3.2.1 Scripts Used by `configure`

`configure` uses some other scripts to help in its work:

- The standard GNU `config.sub` and `config.guess` files, kept in the top level directory, are used.
- The file `config.gcc` is used to handle configuration specific to the particular target machine. The file `config.build` is used to handle configuration specific to the particular build machine. The file `config.host` is used to handle configuration specific to the particular host machine. (In general, these should only be used for features that cannot reasonably be tested in Autoconf feature tests.) See Section 6.3.2.2 [The `config.build`; `config.host`; and `config.gcc` Files], page 64, for details of the contents of these files.
- Each language subdirectory has a file `language/config-lang.in` that is used for front-end-specific configuration. See Section 6.3.8.2 [The Front End `config-lang.in` File], page 73, for details of this file.
- A helper script `configure.frag` is used as part of creating the output of `configure`.

6.3.2.2 The `config.build`; `config.host`; and `config.gcc` Files

The `config.build` file contains specific rules for particular systems which GCC is built on. This should be used as rarely as possible, as the behavior of the build system can always be detected by autoconf.

The `config.host` file contains specific rules for particular systems which GCC will run on. This is rarely needed.

The `config.gcc` file contains specific rules for particular systems which GCC will generate code for. This is usually needed.

Each file has a list of the shell variables it sets, with descriptions, at the top of the file.

FIXME: document the contents of these files, and what variables should be set to control build, host and target configuration.
6.3.2.3 Files Created by configure

Here we spell out what files will be set up by ‘configure’ in the ‘gcc’ directory. Some other files are created as temporary files in the configuration process, and are not used in the subsequent build; these are not documented.

- ‘Makefile’ is constructed from ‘Makefile.in’, together with the host and target fragments (see Chapter 19 [Makefile Fragments], page 563) ‘t-target’ and ‘x-host’ from ‘config’, if any, and language Makefile fragments ‘language/Make-lang.in’.
- ‘auto-host.h’ contains information about the host machine determined by ‘configure’. If the host machine is different from the build machine, then ‘auto-build.h’ is also created, containing such information about the build machine.
- ‘config.status’ is a script that may be run to recreate the current configuration.
- ‘configargs.h’ is a header containing details of the arguments passed to ‘configure’ to configure GCC, and of the thread model used.
- ‘cstamp-h’ is used as a timestamp.
- If a language ‘config-lang.in’ file (see Section 6.3.8.2 [The Front End ‘config-lang.in’ File], page 73) sets outputs, then the files listed in outputs there are also generated.

The following configuration headers are created from the Makefile, using ‘mkconfig.sh’, rather than directly by ‘configure’, ‘config.h’, ‘bconfig.h’ and ‘tconfig.h’ all contain the ‘xm-machine.h’ header, if any, appropriate to the host, build and target machines respectively, the configuration headers for the target, and some definitions; for the host and build machines, these include the autoconfigured headers generated by ‘configure’. The other configuration headers are determined by ‘config.gcc’. They also contain the typedefs for rtx, rtvec and tree.

- ‘config.h’, for use in programs that run on the host machine.
- ‘bconfig.h’, for use in programs that run on the build machine.
- ‘tconfig.h’, for use in programs and libraries for the target machine.
- ‘tm_p.h’, which includes the header ‘machine-protos.h’ that contains prototypes for functions in the target ‘.c’ file. FIXME: why is such a separate header necessary?

6.3.3 Build System in the ‘gcc’ Directory

FIXME: describe the build system, including what is built in what stages. Also list the various source files that are used in the build process but aren’t source files of GCC itself and so aren’t documented below (see Chapter 9 [Passes], page 111).

6.3.4 Makefile Targets

These targets are available from the ‘gcc’ directory:

- **all**: This is the default target. Depending on what your build/host/target configuration is, it coordinates all the things that need to be built.
- **doc**: Produce info-formatted documentation and man pages. Essentially it calls ‘make man’ and ‘make info’.
- **dvi**: Produce DVI-formatted documentation.
pdf  Produce PDF-formatted documentation.
html Produce HTML-formatted documentation.
man  Generate man pages.
info Generate info-formatted pages.
mostlyclean Delete the files made while building the compiler.
clean That, and all the other files built by `make all`.
distclean That, and all the files created by `configure`.
maintainer-clean Distclean plus any file that can be generated from other files. Note that additional tools may be required beyond what is normally needed to build GCC.
srceextra Generates files in the source directory that are not version-controlled but should go into a release tarball.
srcinfo srcman Copies the info-formatted and manpage documentation into the source directory usually for the purpose of generating a release tarball.
install Installs GCC.
uninstall Deletes installed files, though this is not supported.
check Run the testsuite. This creates a `testsuite` subdirectory that has various `.sum` and `.log` files containing the results of the testing. You can run subsets with, for example, `make check-gcc`. You can specify specific tests by setting `RUNTESTFLAGS` to be the name of the `.exp` file, optionally followed by (for some tests) an equals and a file wildcard, like:

```
make check-gcc RUNTESTFLAGS="execute.exp=19980413-**"
```

Note that running the testsuite may require additional tools be installed, such as Tcl or DejaGnu.

The toplevel tree from which you start GCC compilation is not the GCC directory, but rather a complex Makefile that coordinates the various steps of the build, including bootstrapping the compiler and using the new compiler to build target libraries.

When GCC is configured for a native configuration, the default action for `make` is to do a full three-stage bootstrap. This means that GCC is built three times—once with the native compiler, once with the native-built compiler it just built, and once with the compiler it built the second time. In theory, the last two should produce the same results, which `make compare` can check. Each stage is configured separately and compiled into a separate directory, to minimize problems due to ABI incompatibilities between the native compiler and GCC.

If you do a change, rebuilding will also start from the first stage and “bubble” up the change through the three stages. Each stage is taken from its build directory (if it had
been built previously), rebuilt, and copied to its subdirectory. This will allow you to, for example, continue a bootstrap after fixing a bug which causes the stage2 build to crash. It does not provide as good coverage of the compiler as bootstrapping from scratch, but it ensures that the new code is syntactically correct (e.g., that you did not use GCC extensions by mistake), and avoids spurious bootstrap comparison failures\(^1\).

Other targets available from the top level include:

**bootstrap-lean**
Like bootstrap, except that the various stages are removed once they’re no longer needed. This saves disk space.

**bootstrap2**

**bootstrap2-lean**
Performs only the first two stages of bootstrap. Unlike a three-stage bootstrap, this does not perform a comparison to test that the compiler is running properly. Note that the disk space required by a “lean” bootstrap is approximately independent of the number of stages.

**stage\(^N\)-bubble (\(N = 1...4\), profile, feedback)**
Rebuild all the stages up to \(N\), with the appropriate flags, “bubbling” the changes as described above.

**all-stage\(^N\) (\(N = 1...4\), profile, feedback)**
Assuming that stage \(N\) has already been built, rebuild it with the appropriate flags. This is rarely needed.

**cleanstrap**
Remove everything (‘make clean’) and rebuilds (‘make bootstrap’).

**compare**
Compares the results of stages 2 and 3. This ensures that the compiler is running properly, since it should produce the same object files regardless of how it itself was compiled.

**profiledbootstrap**
Builds a compiler with profiling feedback information. In this case, the second and third stages are named ‘profile’ and ‘feedback’, respectively. For more information, see Section “Building with profile feedback” in *Installing GCC*.

**restrap**
Restart a bootstrap, so that everything that was not built with the system compiler is rebuilt.

**stage\(^N\)-start (\(N = 1...4\), profile, feedback)**
For each package that is bootstrapped, rename directories so that, for example, ‘gcc’ points to the stage\(^N\) GCC, compiled with the stage\(N-1\) GCC\(^2\).

You will invoke this target if you need to test or debug the stage\(^N\) GCC. If you only need to execute GCC (but you need not run ‘make’ either to rebuild it or to run test suites), you should be able to work directly in the ‘stage\(^N\)-gcc’ directory. This makes it easier to debug multiple stages in parallel.

---

\(^1\) Except if the compiler was buggy and miscompiled some of the files that were not modified. In this case, it’s best to use `make restrap`.

\(^2\) Customarily, the system compiler is also termed the ‘stage0’ GCC.
For each package that is bootstrapped, relocate its build directory to indicate its stage. For example, if the ‘gcc’ directory points to the stage2 GCC, after invoking this target it will be renamed to ‘stage2-gcc’.

If you wish to use non-default GCC flags when compiling the stage2 and stage3 compilers, set `BOOT_CFLAGS` on the command line when doing `make`.

Usually, the first stage only builds the languages that the compiler is written in: typically, C and maybe Ada. If you are debugging a miscompilation of a different stage2 front-end (for example, of the Fortran front-end), you may want to have front-ends for other languages in the first stage as well. To do so, set `STAGE1_LANGUAGES` on the command line when doing `make`.

For example, in the aforementioned scenario of debugging a Fortran front-end miscompilation caused by the stage1 compiler, you may need a command like

```
make stage2-bubble STAGE1_LANGUAGES=c,fortran
```

Alternatively, you can use per-language targets to build and test languages that are not enabled by default in stage1. For example, `make f95` will build a Fortran compiler even in the stage1 build directory.

### 6.3.5 Library Source Files and Headers under the ‘gcc’ Directory

FIXME: list here, with explanation, all the C source files and headers under the ‘gcc’ directory that aren’t built into the GCC executable but rather are part of runtime libraries and object files, such as ‘crtstuff.c’ and ‘unwind-dw2.c’. See Section 6.3.6 [Headers Installed by GCC], page 68, for more information about the ‘ginclude’ directory.

### 6.3.6 Headers Installed by GCC

In general, GCC expects the system C library to provide most of the headers to be used with it. However, GCC will fix those headers if necessary to make them work with GCC, and will install some headers required of freestanding implementations. These headers are installed in ‘libsubdir/include’. Headers for non-C runtime libraries are also installed by GCC; these are not documented here. (FIXME: document them somewhere.)

Several of the headers GCC installs are in the ‘ginclude’ directory. These headers, ‘iso646.h’, ‘stdarg.h’, ‘stdbool.h’, and ‘stddef.h’, are installed in ‘libsubdir/include’, unless the target Makefile fragment (see Section 19.1 [Target Fragment], page 563) overrides this by setting `USER_H`.

In addition to these headers and those generated by fixing system headers to work with GCC, some other headers may also be installed in ‘libsubdir/include’. ‘config.gcc’ may set `extra_headers`; this specifies additional headers under ‘config’ to be installed on some systems.

GCC installs its own version of `<float.h>`, from ‘ginclude/float.h’. This is done to cope with command-line options that change the representation of floating point numbers.

GCC also installs its own version of `<limits.h>`; this is generated from ‘glimits.h’, together with ‘limitx.h’ and ‘limity.h’ if the system also has its own version of `<limits.h>`. (GCC provides its own header because it is required of ISO C freestanding implementations, but needs to include the system header from its own header as well because other standards such as POSIX specify additional values to be defined in `<limits.h>`.)
Chapter 6: Source Tree Structure and Build System

<limits.h> header is used via ‘libsubdir/include/syslimits.h’, which is copied from ‘gsyslimits.h’ if it does not need fixing to work with GCC; if it needs fixing, ‘syslimits.h’ is the fixed copy.

GCC can also install <tgmath.h>. It will do this when ‘config.gcc’ sets use_gcc_tgmath to yes.

6.3.7 Building Documentation

The main GCC documentation is in the form of manuals in Texinfo format. These are installed in Info format; DVI versions may be generated by ‘make dvi’, PDF versions by ‘make pdf’, and HTML versions by ‘make html’. In addition, some man pages are generated from the Texinfo manuals, there are some other text files with miscellaneous documentation, and runtime libraries have their own documentation outside the ‘gcc’ directory. FIXME: document the documentation for runtime libraries somewhere.

6.3.7.1 Texinfo Manuals

The manuals for GCC as a whole, and the C and C++ front ends, are in files ‘doc/*.texi’. Other front ends have their own manuals in files ‘language/*.texi’. Common files ‘doc/include/*.texi’ are provided which may be included in multiple manuals; the following files are in ‘doc/include’:

‘fdl.texi’
The GNU Free Documentation License.

‘funding.texi’
The section “Funding Free Software”.

‘gcc-common.texi’
Common definitions for manuals.

‘gpl.texi’
‘gpl_v3.texi’
The GNU General Public License.

‘texinfo.tex’
A copy of ‘texinfo.tex’ known to work with the GCC manuals.

DVI-formatted manuals are generated by ‘make dvi’, which uses texi2dvi (via the Makefile macro $(TEXI2DVI)). PDF-formatted manuals are generated by ‘make pdf’, which uses texi2pdf (via the Makefile macro $(TEXI2PDF)). HTML formatted manuals are generated by ‘make html’. Info manuals are generated by ‘make info’ (which is run as part of a bootstrap); this generates the manuals in the source directory, using makeinfo via the Makefile macro $(MAKEINFO), and they are included in release distributions.

Manuals are also provided on the GCC web site, in both HTML and PostScript forms. This is done via the script ‘maintainer-scripts/update_web_docs_svn’. Each manual to be provided online must be listed in the definition of MANUALS in that file; a file ‘name.texi’ must only appear once in the source tree, and the output manual must have the same name as the source file. (However, other Texinfo files, included in manuals but not themselves the root files of manuals, may have names that appear more than once in the source tree.) The manual file ‘name.texi’ should only include other files in its own directory or in
`doc/include`. HTML manuals will be generated by `makeinfo --html`, PostScript manuals by `texi2dvi` and `dvips`, and PDF manuals by `texi2pdf`. All Texinfo files that are parts of manuals must be version-controlled, even if they are generated files, for the generation of online manuals to work.

The installation manual, `doc/install.texi`, is also provided on the GCC web site. The HTML version is generated by the script `doc/install.texi2html`.

### 6.3.7.2 Man Page Generation

Because of user demand, in addition to full Texinfo manuals, man pages are provided which contain extracts from those manuals. These man pages are generated from the Texinfo manuals using `contrib/texi2pod.pl` and `pod2man`. (The man page for `g++`, `cp/g++.1`, just contains a `.so` reference to `gcc.1`, but all the other man pages are generated from Texinfo manuals.)

Because many systems may not have the necessary tools installed to generate the man pages, they are only generated if the `configure` script detects that recent enough tools are installed, and the Makefiles allow generating man pages to fail without aborting the build. Man pages are also included in release distributions. They are generated in the source directory.

Magic comments in Texinfo files starting `@c man` control what parts of a Texinfo file go into a man page. Only a subset of Texinfo is supported by `texi2pod.pl`, and it may be necessary to add support for more Texinfo features to this script when generating new man pages. To improve the man page output, some special Texinfo macros are provided in `doc/include/gcc-common.texi` which `texi2pod.pl` understands:

- `@gcctabopt`
  
  Use in the form `@table @gcctabopt` for tables of options, where for printed output the effect of `@code` is better than that of `@option` but for man page output a different effect is wanted.

- `@gccoptlist`
  
  Use for summary lists of options in manuals.

- `@gol`
  
  Use at the end of each line inside `@gccoptlist`. This is necessary to avoid problems with differences in how the `@gccoptlist` macro is handled by different Texinfo formatters.

FIXME: describe the `texi2pod.pl` input language and magic comments in more detail.

### 6.3.7.3 Miscellaneous Documentation

In addition to the formal documentation that is installed by GCC, there are several other text files in the `gcc` subdirectory with miscellaneous documentation:

- `ABOUT-GCC-NLS`
  
  Notes on GCC’s Native Language Support. FIXME: this should be part of this manual rather than a separate file.

- `ABOUT-NLS`
  
  Notes on the Free Translation Project.
Chapter 6: Source Tree Structure and Build System

'COPYING'
'COPYING3'
The GNU General Public License, Versions 2 and 3.

'COPYING.LIB'
'COPYING3.LIB'
The GNU Lesser General Public License, Versions 2.1 and 3.

'*ChangeLog*'
'*/ChangeLog*'
Change log files for various parts of GCC.

'LANGUAGES'
Details of a few changes to the GCC front-end interface. FIXME: the information in this file should be part of general documentation of the front-end interface in this manual.

'ONEWS'
Information about new features in old versions of GCC. (For recent versions, the information is on the GCC web site.)

'README.Portability'
Information about portability issues when writing code in GCC. FIXME: why isn't this part of this manual or of the GCC Coding Conventions?

FIXME: document such files in subdirectories, at least 'config', 'cp', 'objc', 'testsuite'.

6.3.8 Anatomy of a Language Front End

A front end for a language in GCC has the following parts:

- A directory 'language' under 'gcc' containing source files for that front end. See Section 6.3.8.1 [The Front End 'language' Directory], page 72, for details.
- A mention of the language in the list of supported languages in 'gcc/doc/install.texi'.
- A mention of the name under which the language's runtime library is recognized by '---enable-shared=package' in the documentation of that option in 'gcc/doc/install.texi'.
- A mention of any special prerequisites for building the front end in the documentation of prerequisites in 'gcc/doc/install.texi'.
- Details of contributors to that front end in 'gcc/doc/contrib.texi'. If the details are in that front end's own manual then there should be a link to that manual's list in 'contrib.texi'.
- Information about support for that language in 'gcc/doc/frontends.texi'.
- Information about standards for that language, and the front end’s support for them, in 'gcc/doc/standards.texi'. This may be a link to such information in the front end’s own manual.
- Details of source file suffixes for that language and '-x lang' options supported, in 'gcc/doc/invoke.texi'.
- Entries in default_compilers in 'gcc.c' for source file suffixes for that language.
- Preferably testsuites, which may be under 'gcc/testsuite' or runtime library directories. FIXME: document somewhere how to write testsuite harnesses.
• Probably a runtime library for the language, outside the ‘gcc’ directory. FIXME: document this further.
• Details of the directories of any runtime libraries in ‘gcc/doc/sourcebuild.texi’.
• Check targets in ‘Makefile.def’ for the top-level ‘Makefile’ to check just the compiler or the compiler and runtime library for the language.

If the front end is added to the official GCC source repository, the following are also necessary:
• At least one Bugzilla component for bugs in that front end and runtime libraries. This category needs to be added to the Bugzilla database.
• Normally, one or more maintainers of that front end listed in ‘MAINTAINERS’.
• Mentions on the GCC web site in ‘index.html’ and ‘frontends.html’, with any relevant links on ‘readings.html’. (Front ends that are not an official part of GCC may also be listed on ‘frontends.html’, with relevant links.)
• A news item on ‘index.html’, and possibly an announcement on the gcc-announce@gcc.gnu.org mailing list.
• The front end’s manuals should be mentioned in ‘maintainer-scripts/update_web_docs_svn’ (see Section 6.3.7.1 [Texinfo Manuals], page 69) and the online manuals should be linked to from ‘onlinedocs/index.html’.
• Any old releases or CVS repositories of the front end, before its inclusion in GCC, should be made available on the GCC FTP site ftp://gcc.gnu.org/pub/gcc/old-releases/.
• The release and snapshot script ‘maintainer-scripts/gcc_release’ should be updated to generate appropriate tarballs for this front end.
• If this front end includes its own version files that include the current date, ‘maintainer-scripts/update_version’ should be updated accordingly.

6.3.8.1 The Front End ‘language’ Directory

A front end ‘language’ directory contains the source files of that front end (but not of any runtime libraries, which should be outside the ‘gcc’ directory). This includes documentation, and possibly some subsidiary programs built alongside the front end. Certain files are special and other parts of the compiler depend on their names:

‘config-lang.in’
This file is required in all language subdirectories. See Section 6.3.8.2 [The Front End ‘config-lang.in’ File], page 73, for details of its contents

‘Make-lang.in’
This file is required in all language subdirectories. See Section 6.3.8.3 [The Front End ‘Make-lang.in’ File], page 74, for details of its contents.

‘lang.opt’
This file registers the set of switches that the front end accepts on the command line, and their ‘--help’ text. See Chapter 8 [Options], page 103.

‘lang-specs.h’
This file provides entries for default_compilers in ‘gcc.c’ which override the default of giving an error that a compiler for that language is not installed.
Chapter 6: Source Tree Structure and Build System

'language-tree.def'
This file, which need not exist, defines any language-specific tree codes.

6.3.8.2 The Front End 'config-lang.in' File

Each language subdirectory contains a 'config-lang.in' file. In addition the main directory contains 'c-config-lang.in', which contains limited information for the C language. This file is a shell script that may define some variables describing the language:

language  This definition must be present, and gives the name of the language for some purposes such as arguments to '--enable-languages'.

lang_requires  If defined, this variable lists (space-separated) language front ends other than C that this front end requires to be enabled (with the names given being their language settings). For example, the Java front end depends on the C++ front end, so sets 'lang_requires=c++'.

subdir_requires  If defined, this variable lists (space-separated) front end directories other than C that this front end requires to be present. For example, the Objective-C++ front end uses source files from the C++ and Objective-C front ends, so sets 'subdir_requires="cp objc"'.

target_libs  If defined, this variable lists (space-separated) targets in the top level 'Makefile' to build the runtime libraries for this language, such as target-libobjc.

lang_dirs  If defined, this variable lists (space-separated) top level directories (parallel to 'gcc'), apart from the runtime libraries, that should not be configured if this front end is not built.

build_by_default  If defined to 'no', this language front end is not built unless enabled in a '--enable-languages' argument. Otherwise, front ends are built by default, subject to any special logic in 'configure.ac' (as is present to disable the Ada front end if the Ada compiler is not already installed).

boot_language  If defined to 'yes', this front end is built in stage 1 of the bootstrap. This is only relevant to front ends written in their own languages.

compilers  If defined, a space-separated list of compiler executables that will be run by the driver. The names here will each end with '\$(exeext)'.

outputs  If defined, a space-separated list of files that should be generated by 'configure' substituting values in them. This mechanism can be used to create a file 'language/Makefile' from 'language/Makefile.in', but this is deprecated, building everything from the single 'gcc/Makefile' is preferred.

gtfiles  If defined, a space-separated list of files that should be scanned by 'gengtype.c' to generate the garbage collection tables and routines for this language. This
excludes the files that are common to all front ends. See Chapter 22 [Type Information], page 571.

6.3.8.3 The Front End ‘Make-lang.in’ File

Each language subdirectory contains a ‘Make-lang.in’ file. It contains targets lang.hook (where lang is the setting of language in ‘config-lang.in’) for the following values of hook, and any other Makefile rules required to build those targets (which may if necessary use other Makefiles specified in outputs in ‘config-lang.in’, although this is deprecated). It also adds any testsuite targets that can use the standard rule in ‘gcc/Makefile.in’ to the variable lang_checks.

`all.cross`

`start.encap`

`rest.encap`

FIXME: exactly what goes in each of these targets?

`tags`
Build an etags ‘TAGS’ file in the language subdirectory in the source tree.

`info`
Build info documentation for the front end, in the build directory. This target is only called by `make bootstrap` if a suitable version of makeinfo is available, so does not need to check for this, and should fail if an error occurs.

`dvi`
Build DVI documentation for the front end, in the build directory. This should be done using $(TEXI2DVI), with appropriate ‘-I’ arguments pointing to directories of included files.

`pdf`
Build PDF documentation for the front end, in the build directory. This should be done using $(TEXI2PDF), with appropriate ‘-I’ arguments pointing to directories of included files.

`html`
Build HTML documentation for the front end, in the build directory.

`man`
Build generated man pages for the front end from Texinfo manuals (see Section 6.3.7.2 [Man Page Generation], page 70), in the build directory. This target is only called if the necessary tools are available, but should ignore errors so as not to stop the build if errors occur; man pages are optional and the tools involved may be installed in a broken way.

`install-common`
Install everything that is part of the front end, apart from the compiler executables listed in compilers in ‘config-lang.in’.

`install-info`
Install info documentation for the front end, if it is present in the source directory. This target should have dependencies on info files that should be installed.

`install-man`
Install man pages for the front end. This target should ignore errors.

`install-plugin`
Install headers needed for plugins.

`srcextra`
Copies its dependencies into the source directory. This generally should be used for generated files such as Bison output files which are not version-controlled,
but should be included in any release tarballs. This target will be executed during a bootstrap if ‘--enable-generated-files-in-srcdir’ was specified as a ‘configure’ option.

srcinfo
srcman
Copies its dependencies into the source directory. These targets will be executed during a bootstrap if ‘--enable-generated-files-in-srcdir’ was specified as a ‘configure’ option.

uninstall
Uninstall files installed by installing the compiler. This is currently documented not to be supported, so the hook need not do anything.

mostlyclean
clean
distclean
maintainer-clean
The language parts of the standard GNU ‘*clean’ targets. See Section “Standard Targets for Users” in *GNU Coding Standards*, for details of the standard targets. For GCC, maintainer-clean should delete all generated files in the source directory that are not version-controlled, but should not delete anything that is.

‘Make-lang.in’ must also define a variable lang_OBJECTS to a list of host object files that are used by that language.

### 6.3.9 Anatomy of a Target Back End

A back end for a target architecture in GCC has the following parts:

- A directory ‘machine’ under ‘gcc/config’, containing a machine description ‘machine.md’ file (see Chapter 16 [Machine Descriptions], page 289), header files ‘machine.h’ and ‘machine-protos.h’ and a source file ‘machine.c’ (see Chapter 17 [Target Description Macros and Functions], page 393), possibly a target Makefile fragment ‘t-machine’ (see Section 19.1 [The Target Makefile Fragment], page 563), and maybe some other files. The names of these files may be changed from the defaults given by explicit specifications in ‘config.gcc’.

- If necessary, a file ‘machine-modes.def’ in the ‘machine’ directory, containing additional machine modes to represent condition codes. See Section 17.16 [Condition Code], page 476, for further details.

- An optional ‘machine.opt’ file in the ‘machine’ directory, containing a list of target-specific options. You can also add other option files using the extra_options variable in ‘config.gcc’. See Chapter 8 [Options], page 103.

- Entries in ‘config.gcc’ (see Section 6.3.2.2 [The ‘config.gcc’ File], page 64) for the systems with this target architecture.

- Documentation in ‘gcc/doc/invoke.texi’ for any command-line options supported by this target (see Section 17.3 [Run-time Target Specification], page 400). This means both entries in the summary table of options and details of the individual options.

- Documentation in ‘gcc/doc/extend.texi’ for any target-specific attributes supported (see Section 17.25 [Defining target-specific uses of __attribute__], page 534), including
where the same attribute is already supported on some targets, which are enumerated in the manual.

- Documentation in `gcc/doc/extend.texi` for any target-specific pragmas supported.
- Documentation in `gcc/doc/md.texi` of any target-specific constraint letters (see Section 16.8.5 [Constraints for Particular Machines], page 308).
- A note in `gcc/doc/contrib.texi` under the person or people who contributed the target support.
- Entries in `gcc/doc/install.texi` for all target triplets supported with this target architecture, giving details of any special notes about installation for this target, or saying that there are no special notes if there are none.
- Possibly other support outside the `gcc` directory for runtime libraries. FIXME: reference docs for this. The `libstdc++` porting manual needs to be installed as info for this to work, or to be a chapter of this manual.

If the back end is added to the official GCC source repository, the following are also necessary:

- An entry for the target architecture in `readings.html` on the GCC web site, with any relevant links.
- Details of the properties of the back end and target architecture in `backends.html` on the GCC web site.
- A news item about the contribution of support for that target architecture, in `index.html` on the GCC web site.
- Normally, one or more maintainers of that target listed in `MAINTAINERS`. Some existing architectures may be unmaintained, but it would be unusual to add support for a target that does not have a maintainer when support is added.
7 Testsuites

GCC contains several testsuites to help maintain compiler quality. Most of the runtime libraries and language front ends in GCC have testsuites. Currently only the C language testsuites are documented here; FIXME: document the others.

7.1 Idioms Used in Testsuite Code

In general, C testcases have a trailing ‘-n.c’, starting with ‘-1.c’, in case other testcases with similar names are added later. If the test is a test of some well-defined feature, it should have a name referring to that feature such as ‘feature-1.c’. If it does not test a well-defined feature but just happens to exercise a bug somewhere in the compiler, and a bug report has been filed for this bug in the GCC bug database, ‘prbug-number-1.c’ is the appropriate form of name. Otherwise (for miscellaneous bugs not filed in the GCC bug database), and previously more generally, test cases are named after the date on which they were added. This allows people to tell at a glance whether a test failure is because of a recently found bug that has not yet been fixed, or whether it may be a regression, but does not give any other information about the bug or where discussion of it may be found. Some other language testsuites follow similar conventions.

In the ‘gcc.dg’ testsuite, it is often necessary to test that an error is indeed a hard error and not just a warning—for example, where it is a constraint violation in the C standard, which must become an error with ‘-pedantic-errors’. The following idiom, where the first line shown is line line of the file and the line that generates the error, is used for this:

```c
/* { dg-bogus "warning" "warning in place of error" } */
/* { dg-error "regexp" "message" { target -*-*-* } line } */
```

It may be necessary to check that an expression is an integer constant expression and has a certain value. To check that \( E \) has value \( V \), an idiom similar to the following is used:

```c
char x[((E) == (V) ? 1 : -1)];
```

In ‘gcc.dg’ tests, __typeof__ is sometimes used to make assertions about the types of expressions. See, for example, ‘gcc.dg/c99-condexpr-1.c’. The more subtle uses depend on the exact rules for the types of conditional expressions in the C standard; see, for example, ‘gcc.dg/c99-intconst-1.c’.

It is useful to be able to test that optimizations are being made properly. This cannot be done in all cases, but it can be done where the optimization will lead to code being optimized away (for example, where flow analysis or alias analysis should show that certain code cannot be called) or to functions not being called because they have been expanded as built-in functions. Such tests go in ‘gcc.c-torture/execute’. Where code should be optimized away, a call to a nonexistent function such as link_failure () may be inserted; a definition

```c
#ifndef __OPTIMIZE__
void
link_failure (void)
{
    abort ();
}
#endif
```
will also be needed so that linking still succeeds when the test is run without optimization. When all calls to a built-in function should have been optimized and no calls to the non-built-in version of the function should remain, that function may be defined as \texttt{static} to call \texttt{abort()} (although redeclaring a function as static may not work on all targets).

All testcases must be portable. Target-specific testcases must have appropriate code to avoid causing failures on unsupported systems; unfortunately, the mechanisms for this differ by directory.

FIXME: discuss non-C testsuites here.

\section*{7.2 Directives used within DejaGnu tests}

\subsection*{7.2.1 Syntax and Descriptions of test directives}

Test directives appear within comments in a test source file and begin with \texttt{dg-}. Some of these are defined within DejaGnu and others are local to the GCC testsuite.

The order in which test directives appear in a test can be important: directives local to GCC sometimes override information used by the DejaGnu directives, which know nothing about the GCC directives, so the DejaGnu directives must precede GCC directives.

Several test directives include selectors (see Section 7.2.2 [Selectors], page 81) which are usually preceded by the keyword \texttt{target} or \texttt{xfail}.

\subsection*{7.2.1.1 Specify how to build the test}

\begin{verbatim}
{ dg-do do-what-keyword [{ target/xfail selector }] }
  do-what-keyword specifies how the test is compiled and whether it is executed.
  It is one of:

  preprocess
    Compile with '-E' to run only the preprocessor.

  compile
    Compile with '-S' to produce an assembly code file.

  assemble
    Compile with '-c' to produce a relocatable object file.

  link
    Compile, assemble, and link to produce an executable file.

  run
    Produce and run an executable file, which is expected to return an exit code of 0.

The default is compile. That can be overridden for a set of tests by redefining \texttt{dg-do-what-default} within the .exp file for those tests.

If the directive includes the optional '{ target selector }' then the test is skipped unless the target system matches the selector.

If \texttt{do-what-keyword} is run and the directive includes the optional '{ xfail selector }' and the selector is met then the test is expected to fail. The \texttt{xfail} clause is ignored for other values of \texttt{do-what-keyword}; those tests can use directive \texttt{dg-xfail-if}.
\end{verbatim}
7.2.1.2 Specify additional compiler options

\{ dg-options options [{ target selector }] \}
This DejaGnu directive provides a list of compiler options, to be used if the
target system matches selector, that replace the default options used for this
set of tests.

\{ dg-add-options feature ... \}
Add any compiler options that are needed to access certain features. This
directive does nothing on targets that enable the features by default, or that
don’t provide them at all. It must come after all dg-options directives. For
supported values of feature see Section 7.2.4 [Add Options], page 91.

7.2.1.3 Modify the test timeout value
The normal timeout limit, in seconds, is found by searching the following in order:

- the value defined by an earlier dg-timeout directive in the test
- variable tool_timeout defined by the set of tests
- gcc_timeout set in the target board
- 300

\{ dg-timeout n [{target selector }] \}
Set the time limit for the compilation and for the execution of the test to the
specified number of seconds.

\{ dg-timeout-factor x [{ target selector }] \}
Multiply the normal time limit for compilation and execution of the test by the
specified floating-point factor.

7.2.1.4 Skip a test for some targets

\{ dg-skip-if comment { selector } [{ include-opts } [{ exclude-opts }]] \}
Arguments include-opts and exclude-opts are lists in which each element is
a string of zero or more GCC options. Skip the test if all of the following
conditions are met:

- the test system is included in selector
- for at least one of the option strings in include-opts, every option from that
  string is in the set of options with which the test would be compiled; use
  "*" for an include-opts list that matches any options; that is the default
  if include-opts is not specified
- for each of the option strings in exclude-opts, at least one option from that
  string is not in the set of options with which the test would be compiled;
  use "" for an empty exclude-opts list; that is the default if exclude-opts
  is not specified

For example, to skip a test if option -Os is present:

/** { dg-skip-if "" { *-*-* } { "-Os" } { "" } } */
To skip a test if both options -O2 and -g are present:

/** { dg-skip-if "" { *-*-* } { "-O2 -g" } { "" } } */
To skip a test if either -O2 or -O3 is present:
To skip a test unless option -Os is present:

```c
/* { dg-skip-if "" { *-*-* } { "-O2" "-03" } { "" } } */
```

To skip a test if either -O2 or -O3 is used with -g but not if -fpic is also present:

```c
/* { dg-skip-if "" { *-*-* } { "*" } { "-Os" } } */
```

```c
{ dg-require-effective-target keyword [{ selector }] }
Skip the test if the test target, including current multilib flags, is not covered by the effective-target keyword. If the directive includes the optional `{ selector }' then the effective-target test is only performed if the target system matches the selector. This directive must appear after any dg-do directive in the test and before any dg-additional-sources directive. See Section 7.2.3 [Effective-Target Keywords], page 82.

{ dg-require-support args }
Skip the test if the target does not provide the required support. These directives must appear after any dg-do directive in the test and before any dg-additional-sources directive. They require at least one argument, which can be an empty string if the specific procedure does not examine the argument. See Section 7.2.5 [Require Support], page 92, for a complete list of these directives.

7.2.1.5 Expect a test to fail for some targets

```c
{ dg-xfail-if comment { selector } [{ include-opts } [{ exclude-opts }]]]
Expect the test to fail if the conditions (which are the same as for dg-skip-if) are met. This does not affect the execute step.
```

```c
{ dg-xfail-run-if comment { selector } [{ include-opts } [{ exclude-opts }]]}
Expect the execute step of a test to fail if the conditions (which are the same as for dg-skip-if) are met.
```

7.2.1.6 Expect the test executable to fail

```c
{ dg-shouldfail comment [{ selector } [{ include-opts } [{ exclude-opts }]]]}
Expect the test executable to return a nonzero exit status if the conditions (which are the same as for dg-skip-if) are met.
```

7.2.1.7 Verify compiler messages

```c
{ dg-error regexp [comment [{ target/xfail selector } [line]]]
This DejaGnu directive appears on a source line that is expected to get an error message, or else specifies the source line associated with the message. If there is no message for that line or if the text of that message is not matched by regexp then the check fails and comment is included in the FAIL message. The check does not look for the string ‘error’ unless it is part of regexp.
```

```c
{ dg-warning regexp [comment [{ target/xfail selector } [line]]]
This DejaGnu directive appears on a source line that is expected to get a warning message, or else specifies the source line associated with the message. If there is no message for that line or if the text of that message is not matched
by regexp then the check fails and comment is included in the FAIL message.
The check does not look for the string ‘warning’ unless it is part of regexp.

{ dg-message regexp [comment [{ target/xfail selector } [line] ]]] }
The line is expected to get a message other than an error or warning. If there is
no message for that line or if the text of that message is not matched by regexp
then the check fails and comment is included in the FAIL message.

{ dg-bogus regexp [comment [{ target/xfail selector } [line] ]]] }
This DejaGnu directive appears on a source line that should not get a message
matching regexp, or else specifies the source line associated with the bogus
message. It is usually used with ‘xfail’ to indicate that the message is a
known problem for a particular set of targets.

{ dg-excess-errors comment [{ target/xfail selector }]}
This DejaGnu directive indicates that the test is expected to fail due to compiler
messages that are not handled by ‘dg-error’, ‘dg-warning’ or ‘dg-bogus’. For
this directive ‘xfail’ has the same effect as ‘target’.

{ dg-prune-output regexp }
Prune messages matching regexp from the test output.

7.2.1.8 Verify output of the test executable

{ dg-output regexp [{ target/xfail selector }]}
This DejaGnu directive compares regexp to the combined output that the test
executable writes to ‘stdout’ and ‘stderr’.

7.2.1.9 Specify additional files for a test

{ dg-additional-files "filelist" }
Specify additional files, other than source files, that must be copied to the
system where the compiler runs.

{ dg-additional-sources "filelist" }
Specify additional source files to appear in the compile line following the main
test file.

7.2.1.10 Add checks at the end of a test

{ dg-final { local-directive } }
This DejaGnu directive is placed within a comment anywhere in the source file
and is processed after the test has been compiled and run. Multiple ‘dg-final’
commands are processed in the order in which they appear in the source file.
See Section 7.2.6 [Final Actions], page 93, for a list of directives that can be
used within dg-final.

7.2.2 Selecting targets to which a test applies
Several test directives include selectors to limit the targets for which a test is run or to
declare that a test is expected to fail on particular targets.

A selector is:
• one or more target triplets, possibly including wildcard characters
• a single effective-target keyword (see Section 7.2.3 [Effective-Target Keywords], page 82)
• a logical expression

Depending on the context, the selector specifies whether a test is skipped and reported as unsupported or is expected to fail. Use ‘*-*-*' to match any target.

A selector expression appears within curly braces and uses a single logical operator: one of ‘!’, ‘&&’, or ‘||’. An operand is another selector expression, an effective-target keyword, a single target triplet, or a list of target triplets within quotes or curly braces. For example:

```
{ target { ! "hppa*-*-* ia64*-*-*" } }
{ target { powerpc*-*-* && lp64 } }
{ xfail { lp64 || vect_no_align } }
```

### 7.2.3 Keywords describing target attributes

Effective-target keywords identify sets of targets that support particular functionality. They are used to limit tests to be run only for particular targets, or to specify that particular sets of targets are expected to fail some tests.

Effective-target keywords are defined in ‘lib/target-supports.exp’ in the GCC test-suite, with the exception of those that are documented as being local to a particular test directory.

The ‘effective target’ takes into account all of the compiler options with which the test will be compiled, including the multilib options. By convention, keywords ending in _nocache can also include options specified for the particular test in an earlier dg-options or dg-add-options directive.

#### 7.2.3.1 Data type sizes

- **ilp32** Target has 32-bit int, long, and pointers.
- **lp64** Target has 32-bit int, 64-bit long and pointers.
- **l1p64** Target has 32-bit int and long, 64-bit long long and pointers.
- **double64** Target has 64-bit double.
- **double64plus** Target has double that is 64 bits or longer.
- **int32plus** Target has int that is at 32 bits or longer.
- **int16** Target has int that is 16 bits or shorter.
- **large_double** Target supports double that is longer than float.
- **large_long_double** Target supports long double that is longer than double.
- **ptr32plus** Target has pointers that are 32 bits or longer.
size32plus
   Target supports array and structure sizes that are 32 bits or longer.

4byte_wchar_t
   Target has wchar_t that is at least 4 bytes.

7.2.3.2 Fortran-specific attributes

fortran_integer_16
   Target supports Fortran integer that is 16 bytes or longer.

fortran_large_int
   Target supports Fortran integer kinds larger than integer(8).

fortran_large_real
   Target supports Fortran real kinds larger than real(8).

7.2.3.3 Vector-specific attributes

vect_condition
   Target supports vector conditional operations.

vect_double
   Target supports hardware vectors of double.

vect_float
   Target supports hardware vectors of float.

vect_int
   Target supports hardware vectors of int.

vect_long
   Target supports hardware vectors of long.

vect_long_long
   Target supports hardware vectors of long long.

vect_aligned_arrays
   Target aligns arrays to vector alignment boundary.

vect_hw_misalign
   Target supports a vector misalign access.

vect_no_align
   Target does not support a vector alignment mechanism.

vect_no_int_max
   Target does not support a vector max instruction on int.

vect_no_int_add
   Target does not support a vector add instruction on int.

vect_no_bitwise
   Target does not support vector bitwise instructions.

vect_char_mult
   Target supports vector char multiplication.
vect_short_mult
    Target supports vector short multiplication.

vect_int_mult
    Target supports vector int multiplication.

vect_extract_even_odd
    Target supports vector even/odd element extraction.

vect_extract_even_odd_wide
    Target supports vector even/odd element extraction of vectors with elements \texttt{S1mode} or larger.

vect_interleave
    Target supports vector interleaving.

vect_strided
    Target supports vector interleaving and extract even/odd.

vect_strided_wide
    Target supports vector interleaving and extract even/odd for wide element types.

vect_perm
    Target supports vector permutation.

vect_shift
    Target supports a hardware vector shift operation.

vect_widen_sum_hi_to_si
    Target supports a vector widening summation of \texttt{short} operands into \texttt{int} results, or can promote (unpack) from \texttt{short} to \texttt{int}.

vect_widen_sum_qi_to_hi
    Target supports a vector widening summation of \texttt{char} operands into \texttt{short} results, or can promote (unpack) from \texttt{char} to \texttt{short}.

vect_widen_sum_qi_to_si
    Target supports a vector widening summation of \texttt{char} operands into \texttt{int} results.

vect_widen_mult_qi_to_hi
    Target supports a vector widening multiplication of \texttt{char} operands into \texttt{short} results, or can promote (unpack) from \texttt{char} to \texttt{short} and perform non-widening multiplication of \texttt{short}.

vect_widen_mult_hi_to_si
    Target supports a vector widening multiplication of \texttt{short} operands into \texttt{int} results, or can promote (unpack) from \texttt{short} to \texttt{int} and perform non-widening multiplication of \texttt{int}.

vect_sdot_qi
    Target supports a vector dot-product of \texttt{signed char}.

vect_udot_qi
    Target supports a vector dot-product of \texttt{unsigned char}.
vect_sdot_hi
  Target supports a vector dot-product of signed short.

vect_udot_hi
  Target supports a vector dot-product of unsigned short.

vect_pack_trunc
  Target supports a vector demotion (packing) of short to char and from int to short using modulo arithmetic.

vect_unpack
  Target supports a vector promotion (unpacking) of char to short and from char to int.

vect_intfloat_cvt
  Target supports conversion from signed int to float.

vect_uintfloat_cvt
  Target supports conversion from unsigned int to float.

vect_floatint_cvt
  Target supports conversion from float to signed int.

vect_floatuint_cvt
  Target supports conversion from float to unsigned int.

7.2.3.4 Thread Local Storage attributes

tls
  Target supports thread-local storage.

tls_native
  Target supports native (rather than emulated) thread-local storage.

tls_runtime
  Test system supports executing TLS executables.

7.2.3.5 Decimal floating point attributes

dfp
  Target supports compiling decimal floating point extension to C.

dfp_nocache
  Including the options used to compile this particular test, the target supports compiling decimal floating point extension to C.

dfp_nocache
  Test system can execute decimal floating point tests.

dfp_nocache
  Including the options used to compile this particular test, the test system can execute decimal floating point tests.

hard_dfp
  Target generates decimal floating point instructions with current options.

7.2.3.6 ARM-specific attributes

arm32
  ARM target generates 32-bit code.

arm_eabi
  ARM target adheres to the ABI for the ARM Architecture.
arm_hard_vfp_ok
    ARM target supports -mfpu=vfp -mfloat-abi=hard. Some multilibs may be incompatible with these options.

arm_iwmmxt_ok
    ARM target supports -mcpu=iwmmxt. Some multilibs may be incompatible with this option.

arm_neon
    ARM target supports generating NEON instructions.

arm_neon_hw
    Test system supports executing NEON instructions.

arm_neon_ok
    ARM Target supports -mfpu=neon -mfloat-abi=softfp or compatible options. Some multilibs may be incompatible with these options.

arm_neon_fp16_ok
    ARM Target supports -mfpu=neon-fp16 -mfloat-abi=softfp or compatible options. Some multilibs may be incompatible with these options.

arm_thumb1_ok
    ARM target generates Thumb-1 code for -mthumb.

arm_thumb2_ok
    ARM target generates Thumb-2 code for -mthumb.

arm_vfp_ok
    ARM target supports -mfpu=vfp -mfloat-abi=softfp. Some multilibs may be incompatible with these options.

7.2.3.7 MIPS-specific attributes

mips64
    MIPS target supports 64-bit instructions.

nomips16
    MIPS target does not produce MIPS16 code.

mips16_attribute
    MIPS target can generate MIPS16 code.

mips_loongson
    MIPS target is a Loongson-2E or -2F target using an ABI that supports the Loongson vector modes.

mips_newabi_large_long_double
    MIPS target supports long double larger than double when using the new ABI.

mpaired_single
    MIPS target supports -mpaired-single.

7.2.3.8 PowerPC-specific attributes

powerpc64
    Test system supports executing 64-bit instructions.
powerpc_altivec
   PowerPC target supports AltiVec.

powerpc_altivec_ok
   PowerPC target supports -maltivec.

powerpc_fprs
   PowerPC target supports floating-point registers.

powerpc_hard_double
   PowerPC target supports hardware double-precision floating-point.

powerpc_ppu_ok
   PowerPC target supports -mcpu=cell.

powerpc_spe
   PowerPC target supports PowerPC SPE.

powerpc_spe_nocache
   Including the options used to compile this particular test, the PowerPC target supports PowerPC SPE.

powerpc_spu
   PowerPC target supports PowerPC SPU.

spu_auto_overlay
   SPU target has toolchain that supports automatic overlay generation.

powerpc_vsx_ok
   PowerPC target supports -mvsx.

powerpc_405_nocache
   Including the options used to compile this particular test, the PowerPC target supports PowerPC 405.

vmx_hw
   PowerPC target supports executing AltiVec instructions.

7.2.3.9 Other hardware attributes

avx
   Target supports compiling avx instructions.

avx_runtime
   Target supports the execution of avx instructions.

cell_hw
   Test system can execute AltiVec and Cell PPU instructions.

coldfire_fpu
   Target uses a ColdFire FPU.

hard_float
   Target supports FPU instructions.

sse
   Target supports compiling sse instructions.

sse_runtime
   Target supports the execution of sse instructions.

sse2
   Target supports compiling sse2 instructions.
sse2_runtime
Target supports the execution of sse2 instructions.

sync_char_short
Target supports atomic operations on char and short.

sync_int_long
Target supports atomic operations on int and long.

ultrasparc_hw
Test environment appears to run executables on a simulator that accepts only EM_SPARC executables and chokes on EM_SPARC32PLUS or EM_SPARCV9 executables.

dummywcsftime
Target uses a dummy _wcsftime function that always returns zero.

fd_truncate
Target can truncate a file from a file descriptor, as used by ‘libgfortran/io/unix.c:fd_truncate’; i.e. ftruncate or chsize.

freestanding
Target is ‘freestanding’ as defined in section 4 of the C99 standard. Effectively, it is a target which supports no extra headers or libraries other than what is considered essential.

init_priority
Target supports constructors with initialization priority arguments.

inttypes_types
Target has the basic signed and unsigned types in _inttypes.h. This is for tests that GCC’s notions of these types agree with those in the header, as some systems have only _inttypes.h.

lax_strtofp
Target might have errors of a few ULP in string to floating-point conversion functions and overflow is not always detected correctly by those functions.

newlib
Target supports Newlib.
Chapter 7: Testsuites

pow10  Target provides pow10 function.

pthread  Target can compile using pthread.h with no errors or warnings.

pthread_h  Target has pthread.h.

run_expensive_tests  Expensive testcases (usually those that consume excessive amounts of CPU time) should be run on this target. This can be enabled by setting the GCC_TEST_RUN_EXPENSIVE environment variable to a non-empty string.

simulator  Test system runs executables on a simulator (i.e. slowly) rather than hardware (i.e. fast).

stdint_types  Target has the basic signed and unsigned C types in stdint.h. This will be obsolete when GCC ensures a working stdint.h for all targets.

trampolines  Target supports trampolines.

uclibc  Target supports uClibc.

unwrapped  Target does not use a status wrapper.

vxworks_kernel  Target is a VxWorks kernel.

vxworks_rtp  Target is a VxWorks RTP.

wchar  Target supports wide characters.

7.2.3.11 Other attributes

automatic_stack_alignment  Target supports automatic stack alignment.

cxa_atexit  Target uses __cxa_atexit.

default_packed  Target has packed layout of structure members by default.

fgraphite  Target supports Graphite optimizations.

fixed_point  Target supports fixed-point extension to C.

fopenmp  Target supports OpenMP via ‘-fopenmp’.

fpic  Target supports ‘-fpic’ and ‘-fPIC’.

freorder  Target supports ‘-freorder-blocks-and-partition’.
fstack_protector
Target supports ‘-fstack-protector’.

gas
Target uses GNU as.

gc_sections
Target supports ‘--gc-sections’.

keeps_null_pointer_checks
Target keeps null pointer checks, either due to the use of
‘-fno-delete-null-pointer-checks’ or hardwired into the target.

tlto
Compiler has been configured to support link-time optimization (LTO).

named_sections
Target supports named sections.

natural_alignment_32
Target uses natural alignment (aligned to type size) for types of 32 bits or less.

target_natural_alignment_64
Target uses natural alignment (aligned to type size) for types of 64 bits or less.

nonpic
Target does not generate PIC by default.

pcc_bitfield_type_matters
Target defines PCC_BITFIELD_TYPE_MATTERS.

pe_aligned_commons
Target supports ‘-mpe-aligned-commons’.

section_anchors
Target supports section anchors.

short Enums
Target defaults to short enums.

static
Target supports ‘-static’.

static_libgfortran
Target supports statically linking ‘libgfortran’.

string_merging
Target supports merging string constants at link time.

ucn
Target supports compiling and assembling UCN.

ucn_nocache
Including the options used to compile this particular test, the target supports
compiling and assembling UCN.

unaligned_stack
Target does not guarantee that its STACK_BOUNDARY is greater than or equal to
the required vector alignment.

vector_alignment_reachable
Vector alignment is reachable for types of 32 bits or less.
vector_alignment_reachable_for_64bit
Vector alignment is reachable for types of 64 bits or less.

wchar_t_char16_t_compatible
Target supports wchar_t that is compatible with char16_t.

wchar_t_char32_t_compatible
Target supports wchar_t that is compatible with char32_t.

7.2.3.12 Local to tests in gcc.target/i386

3dnow Target supports compiling 3dnow instructions.
aes Target supports compiling aes instructions.
fma4 Target supports compiling fma4 instructions.
ms_hook_prologue Target supports attribute ms_hook_prologue.
pclmul Target supports compiling pclmul instructions.
sse3 Target supports compiling sse3 instructions.
sse4 Target supports compiling sse4 instructions.
sse4a Target supports compiling sse4a instructions.
ssse3 Target supports compiling ssse3 instructions.
vaes Target supports compiling vaes instructions.
vpclmul Target supports compiling vpclmul instructions.
xop Target supports compiling xop instructions.

7.2.3.13 Local to tests in gcc.target/spu/ea
ealib Target __ea library functions are available.

7.2.3.14 Local to tests in gcc.test-framework

no Always returns 0.

yes Always returns 1.

7.2.4 Features for dg-add-options
The supported values of feature for directive dg-add-options are:

arm_neon NEON support. Only ARM targets support this feature, and only then in certain modes; see the [arm_neon_ok effective target keyword], page 86.

arm_neon_fp16 NEON and half-precision floating point support. Only ARM targets support this feature, and only then in certain modes; see the [arm_neon_fp16_ok effective target keyword], page 86.
bind_pic_locally
Add the target-specific flags needed to enable functions to bind locally when using pic/PIC passes in the testsuite.

c99_runtime
Add the target-specific flags needed to access the C99 runtime.

ieee
Add the target-specific flags needed to enable full IEEE compliance mode.

mips16_attribute
mips16 function attributes. Only MIPS targets support this feature, and only then in certain modes.

tls
Add the target-specific flags needed to use thread-local storage.

7.2.5 Variants of dg-require-support
A few of the dg-require directives take arguments.

dg-require-icnv codeset
Skip the test if the target does not support icnv. codeset is the codeset to convert to.

dg-require-profiling profopt
Skip the test if the target does not support profiling with option profopt.

dg-require-visibility vis
Skip the test if the target does not support the visibility attribute. If vis is "", support for visibility("hidden") is checked, for visibility("vis") otherwise.

The original dg-require directives were defined before there was support for effective-target keywords. The directives that do not take arguments could be replaced with effective-target keywords.

dg-require-alias ""
Skip the test if the target does not support the ‘alias’ attribute.

dg-require-ascii-locale ""
Skip the test if the host does not support an ASCII locale.

dg-require-compat-dfp ""
Skip this test unless both compilers in a ‘compat’ testsuite support decimal floating point.

dg-require-cxa-atexit ""
Skip the test if the target does not support __cxa_atexit. This is equivalent to dg-require-effective-target cxa_atexit.

dg-require-dll ""
Skip the test if the target does not support DLL attributes.

dg-require-fork ""
Skip the test if the target does not support fork.
dg-require-gc-sections ""
Skip the test if the target’s linker does not support the --gc-sections flags. This is equivalent to dg-require-effective-target gc-sections.

dg-require-host-local ""
Skip the test if the host is remote, rather than the same as the build system. Some tests are incompatible with DejaGnu’s handling of remote hosts, which involves copying the source file to the host and compiling it with a relative path and "-o a.out".

dg-require-mkfifo ""
Skip the test if the target does not support mkfifo.

dg-require-named-sections ""
Skip the test if the target does not support named sections. This is equivalent to dg-require-effective-target named_sections.

dg-require-weak ""
Skip the test if the target does not support weak symbols.

dg-require-weak-overwrite ""
Skip the test if the target does not support overriding weak symbols.

7.2.6 Commands for use in dg-final
The GCC testsuite defines the following directives to be used within dg-final.

7.2.6.1 Scan a particular file

  scan-file filename regexp [{ target/xfail selector }]
  Passes if regexp matches text in filename.

  scan-file-not filename regexp [{ target/xfail selector }]
  Passes if regexp does not match text in filename.

  scan-module module regexp [{ target/xfail selector }]
  Passes if regexp matches in Fortran module module.

7.2.6.2 Scan the assembly output

  scan-assembler regex [{ target/xfail selector }]
  Passes if regex matches text in the test’s assembler output.

  scan-assembler-not regex [{ target/xfail selector }]
  Passes if regex does not match text in the test’s assembler output.

  scan-assembler-times regex num [{ target/xfail selector }]
  Passes if regex is matched exactly num times in the test’s assembler output.

  scan-assembler-dem regex [{ target/xfail selector }]
  Passes if regex matches text in the test’s demangled assembler output.

  scan-assembler-dem-not regex [{ target/xfail selector }]
  Passes if regex does not match text in the test’s demangled assembler output.
scan-hidden symbol [{ target/xfail selector }]
   Passes if symbol is defined as a hidden symbol in the test’s assembly output.

scan-not-hidden symbol [{ target/xfail selector }]
   Passes if symbol is not defined as a hidden symbol in the test’s assembly output.

### 7.2.6.3 Scan optimization dump files
These commands are available for kind of tree, rtl, and ipa.

scan-kind-dump regex suffix [{ target/xfail selector }]
   Passes if regex matches text in the dump file with suffix suffix.

scan-kind-dump-not regex suffix [{ target/xfail selector }]
   Passes if regex does not match text in the dump file with suffix suffix.

scan-kind-dump-times regex num suffix [{ target/xfail selector }]
   Passes if regex is found exactly num times in the dump file with suffix suffix.

scan-kind-dump-dem regex suffix [{ target/xfail selector }]
   Passes if regex matches demangled text in the dump file with suffix suffix.

scan-kind-dump-dem-not regex suffix [{ target/xfail selector }]
   Passes if regex does not match demangled text in the dump file with suffix suffix.

### 7.2.6.4 Verify that an output files exists or not

output-exists [{ target/xfail selector }]
   Passes if compiler output file exists.

output-exists-not [{ target/xfail selector }]
   Passes if compiler output file does not exist.

### 7.2.6.5 Check for LTO tests

scan-symbol regexp [{ target/xfail selector }]
   Passes if the pattern is present in the final executable.

### 7.2.6.6 Checks for gcov tests

run-gcov sourcefile
   Check line counts in gcov tests.

run-gcov [branches] [calls] { opts sourcefile }
   Check branch and/or call counts, in addition to line counts, in gcov tests.

### 7.2.6.7 Clean up generated test files

cleanup-coverage-files
   Removes coverage data files generated for this test.

cleanup-ipa-dump suffix
   Removes IPA dump files generated for this test.

cleanup-modules
   Removes Fortran module files generated for this test.
cleanup-profile-file
Removes profiling files generated for this test.

cleanup-repo-files
Removes files generated for this test for ‘-frepo’.

cleanup-rtl-dump suffix
Removes RTL dump files generated for this test.

cleanup-saved-temps
Removes files for the current test which were kept for ‘-save-temps’.

cleanup-tree-dump suffix
Removes tree dump files matching suffix which were generated for this test.

7.3 Ada Language Testsuites

The Ada testsuite includes executable tests from the ACATS 2.5 testsuite, publicly available at http://www.adaic.org/compilers/acats/2.5.

These tests are integrated in the GCC testsuite in the ‘ada/acats’ directory, and enabled automatically when running make check, assuming the Ada language has been enabled when configuring GCC.

You can also run the Ada testsuite independently, using make check-ada, or run a subset of the tests by specifying which chapter to run, e.g.:

$ make check-ada CHAPTERS="c3 c9"

The tests are organized by directory, each directory corresponding to a chapter of the Ada Reference Manual. So for example, ‘c9’ corresponds to chapter 9, which deals with tasking features of the language.

There is also an extra chapter called ‘gcc’ containing a template for creating new executable tests, although this is deprecated in favor of the ‘gnat.dg’ testsuite.

The tests are run using two sh scripts: ‘run_acats’ and ‘run_all.sh’. To run the tests using a simulator or a cross target, see the small customization section at the top of ‘run_all.sh’.

These tests are run using the build tree: they can be run without doing a make install.

7.4 C Language Testsuites

GCC contains the following C language testsuites, in the ‘gcc/testsuite’ directory:

‘gcc.dg’ This contains tests of particular features of the C compiler, using the more modern ‘dg’ harness. Correctness tests for various compiler features should go here if possible.

Magic comments determine whether the file is preprocessed, compiled, linked or run. In these tests, error and warning message texts are compared against expected texts or regular expressions given in comments. These tests are run with the options ‘-ansi -pedantic’ unless other options are given in the test. Except as noted below they are not run with multiple optimization options.
'gcc.dg/compat'
This subdirectory contains tests for binary compatibility using 'lib/compat.exp', which in turn uses the language-independent support (see Section 7.9 [Support for testing binary compatibility], page 100).

'gcc.dg/cpp'
This subdirectory contains tests of the preprocessor.

'gcc.dg/debug'
This subdirectory contains tests for debug formats. Tests in this subdirectory are run for each debug format that the compiler supports.

'gcc.dg/format'
This subdirectory contains tests of the '-Wformat' format checking. Tests in this directory are run with and without '-DWIDE'.

'gcc.dg/noncompile'
This subdirectory contains tests of code that should not compile and does not need any special compilation options. They are run with multiple optimization options, since sometimes invalid code crashes the compiler with optimization.

'gcc.dg/special'
FIXME: describe this.

'gcc.c-torture'
This contains particular code fragments which have historically broken easily. These tests are run with multiple optimization options, so tests for features which only break at some optimization levels belong here. This also contains tests to check that certain optimizations occur. It might be worthwhile to separate the correctness tests cleanly from the code quality tests, but it hasn’t been done yet.

'gcc.c-torture/compat'
FIXME: describe this.
This directory should probably not be used for new tests.

'gcc.c-torture/compile'
This testsuite contains test cases that should compile, but do not need to link or run. These test cases are compiled with several different combinations of optimization options. All warnings are disabled for these test cases, so this directory is not suitable if you wish to test for the presence or absence of compiler warnings. While special options can be set, and tests disabled on specific platforms, by the use of '.x' files, mostly these test cases should not contain platform dependencies. FIXME: discuss how defines such as NO_LABEL_VALUES and STACK_SIZE are used.

'gcc.c-torture/execute'
This testsuite contains test cases that should compile, link and run; otherwise the same comments as for 'gcc.c-torture/compile' apply.

'gcc.c-torture/execute/ieee'
This contains tests which are specific to IEEE floating point.
'gcc.c-torture/unsorted'

FIXME: describe this.

This directory should probably not be used for new tests.

'gcc.misc-tests'

This directory contains C tests that require special handling. Some of these tests have individual expect files, and others share special-purpose expect files:

'bprob*.c'

Test `-fbranch-probabilities` using `gcc.misc-tests/bprob.exp`, which in turn uses the generic, language-independent framework (see Section 7.8 [Support for testing profile-directed optimizations], page 99).

'gcov*.c'

Test gcov output using `gcov.exp`, which in turn uses the language-independent support (see Section 7.7 [Support for testing gcov], page 98).

'i386-pf-*.c'

Test i386-specific support for data prefetch using `i386-prefetch.exp`.

'gcc.test-framework'

'dg-*.c'

Test the testsuite itself using `gcc.test-framework/test-framework.exp`.

FIXME: merge in `testsuite/README.gcc` and discuss the format of test cases and magic comments more.

7.5 The Java library testsuites.

Runtime tests are executed via `make check` in the `target/libjava/testsuite` directory in the build tree. Additional runtime tests can be checked into this testsuite.

Regression testing of the core packages in libgcj is also covered by the Mauve testsuite. The Mauve Project develops tests for the Java Class Libraries. These tests are run as part of libgcj testing by placing the Mauve tree within the libjava testsuite sources at `libjava/testsuite/libjava.mauve/mauve`, or by specifying the location of that tree when invoking `make`, as in `make MAUVEDIR=~/mauve check`.

To detect regressions, a mechanism in `mauve.exp` compares the failures for a test run against the list of expected failures in `libjava/testsuite/libjava.mauve/xfails` from the source hierarchy. Update this file when adding new failing tests to Mauve, or when fixing bugs in libgcj that had caused Mauve test failures.

We encourage developers to contribute test cases to Mauve.

7.6 Support for testing link-time optimizations

Tests for link-time optimizations usually require multiple source files that are compiled separately, perhaps with different sets of options. There are several special-purpose test directives used for these tests.

{ dg-lto-do do-what-keyword }

`do-what-keyword` specifies how the test is compiled and whether it is executed.

It is one of:
assemble Compile with ‘-c’ to produce a relocatable object file.

link Compile, assemble, and link to produce an executable file.

run Produce and run an executable file, which is expected to return an exit code of 0.

The default is assemble. That can be overridden for a set of tests by redefining \texttt{dg-do-what-default} within the .\texttt{exp} file for those tests.

Unlike \texttt{dg-do}, \texttt{dg-lto-do} does not support an optional ‘\texttt{target}’ or ‘\texttt{xfail}’ list. Use \texttt{dg-skip-if}, \texttt{dg-xfail-if}, or \texttt{dg-xfail-run-if}.

\{ \texttt{dg-lto-options} \{ \{ \texttt{options} \} \{ \{ \texttt{options} \} \} \{ \{ \texttt{target selector} \} \} \}

This directive provides a list of one or more sets of compiler options to override \texttt{LTO\_OPTIONS}. Each test will be compiled and run with each of these sets of options.

\{ \texttt{dg-extra-ld-options} \texttt{options} \{ \{ \texttt{target selector} \} \}

This directive adds \texttt{options} to the linker options used.

\{ \texttt{dg-suppress-ld-options} \texttt{options} \{ \{ \texttt{target selector} \} \}

This directive removes \texttt{options} from the set of linker options used.

\section{Support for testing gcov}

Language-independent support for testing gcov, and for checking that branch profiling produces expected values, is provided by the expect file ‘lib/gcov.exp’. gcov tests also rely on procedures in ‘lib/gcc-dg.exp’ to compile and run the test program. A typical gcov test contains the following DejaGnu commands within comments:

\begin{verbatim}
\{ \texttt{dg-options} "-fprofile-arcs -ftest-coverage" \}
\{ \texttt{dg-do run} \{ \texttt{target native} \} \}
\{ \texttt{dg-final} \{ \texttt{run-gcov sourcefile} \} \}
\end{verbatim}

Checks of gcov output can include line counts, branch percentages, and call return percentages. All of these checks are requested via commands that appear in comments in the test’s source file. Commands to check line counts are processed by default. Commands to check branch percentages and call return percentages are processed if the \texttt{run-gcov} command has arguments \texttt{branches} or \texttt{calls}, respectively. For example, the following specifies checking both, as well as passing ‘-b’ to gcov:

\begin{verbatim}
\{ \texttt{dg-final} \{ \texttt{run-gcov branches calls} \{ -b sourcefile \} \} \}
\end{verbatim}

A line count command appears within a comment on the source line that is expected to get the specified count and has the form \texttt{count(cnt)}. A test should only check line counts for lines that will get the same count for any architecture.

Commands to check branch percentages (\texttt{branch}) and call return percentages (\texttt{returns}) are very similar to each other. A beginning command appears on or before the first of a range of lines that will report the percentage, and the ending command follows that range of lines. The beginning command can include a list of percentages, all of which are expected to be found within the range. A range is terminated by the next command of the same kind. A command \texttt{branch(\texttt{end})} or \texttt{returns(\texttt{end})} marks the end of a range without starting a new one. For example:
if (i > 10 && j > i && j < 20) /* branch(27 50 75) */
    /* branch(end) */
    foo (i, j);

For a call return percentage, the value specified is the percentage of calls reported to
return. For a branch percentage, the value is either the expected percentage or 100 mi-
nus that value, since the direction of a branch can differ depending on the target or the
optimization level.

Not all branches and calls need to be checked. A test should not check for branches that
might be optimized away or replaced with predicated instructions. Don’t check for calls
inserted by the compiler or ones that might be inlined or optimized away.

A single test can check for combinations of line counts, branch percentages, and call
return percentages. The command to check a line count must appear on the line that will
report that count, but commands to check branch percentages and call return percentages
can bracket the lines that report them.

7.8 Support for testing profile-directed optimizations

The file `profopt.exp` provides language-independent support for checking correct execu-
tion of a test built with profile-directed optimization. This testing requires that a test
program be built and executed twice. The first time it is compiled to generate profile data,
and the second time it is compiled to use the data that was generated during the first
execution. The second execution is to verify that the test produces the expected results.

To check that the optimization actually generated better code, a test can be built and
run a third time with normal optimizations to verify that the performance is better with the
profile-directed optimizations. `profopt.exp` has the beginnings of this kind of support.

`profopt.exp` provides generic support for profile-directed optimizations. Each set of
tests that uses it provides information about a specific optimization:

tool tool being tested, e.g., gcc

profile_option options used to generate profile data

feedback_option options used to optimize using that profile data

prof_ext suffix of profile data files

PROFOPT_OPTIONS list of options with which to run each test, similar to the lists for torture tests

{ dg-final-generate { local-directive } }

This directive is similar to dg-final, but the local-directive is run after the
generation of profile data.

{ dg-final-use { local-directive } }

The local-directive is run after the profile data have been used.
7.9 Support for testing binary compatibility

The file ‘compat.exp’ provides language-independent support for binary compatibility testing. It supports testing interoperability of two compilers that follow the same ABI, or of multiple sets of compiler options that should not affect binary compatibility. It is intended to be used for testsuites that complement ABI testsuites.

A test supported by this framework has three parts, each in a separate source file: a main program and two pieces that interact with each other to split up the functionality being tested.

‘testname_main.suffix’
Contains the main program, which calls a function in file ‘testname_x.suffix’.

‘testname_x.suffix’
Contains at least one call to a function in ‘testname_y.suffix’.

‘testname_y.suffix’
Shares data with, or gets arguments from, ‘testname_x.suffix’.

Within each test, the main program and one functional piece are compiled by the GCC under test. The other piece can be compiled by an alternate compiler. If no alternate compiler is specified, then all three source files are all compiled by the GCC under test. You can specify pairs of sets of compiler options. The first element of such a pair specifies options used with the GCC under test, and the second element of the pair specifies options used with the alternate compiler. Each test is compiled with each pair of options.

‘compat.exp’ defines default pairs of compiler options. These can be overridden by defining the environment variable COMPAT_OPTIONS as:

```
COMPAT_OPTIONS="[list [list {tst1} {alti}] ...
...[list {tstm} {altn}]"
```

where tsti and alti are lists of options, with tsti used by the compiler under test and alti used by the alternate compiler. For example, with `list [list {-g -00} {-03}] [list {-fpic} {-fPIC -O2}]`, the test is first built with ‘-g -00’ by the compiler under test and with ‘-03’ by the alternate compiler. The test is built a second time using ‘-fpic’ by the compiler under test and ‘-fPIC -02’ by the alternate compiler.

An alternate compiler is specified by defining an environment variable to be the full pathname of an installed compiler; for C define ALT_CC_UNDER_TEST, and for C++ define ALT_CXX_UNDER_TEST. These will be written to the ‘site.exp’ file used by DejaGnu. The default is to build each test with the compiler under test using the first of each pair of compiler options from COMPAT_OPTIONS. When ALT_CC_UNDER_TEST or ALT_CXX_UNDER_TEST is same, each test is built using the compiler under test but with combinations of the options from COMPAT_OPTIONS.

To run only the C++ compatibility suite using the compiler under test and another version of GCC using specific compiler options, do the following from ‘objdir/gcc’:

```
rmdir site.exp
make -k \
    ALT_CXX_UNDER_TEST=${alt_prefix}/bin/g++ \n    COMPAT_OPTIONS="lists as shown above" \n    check-c++ \
    RUNTESTFLAGS="compat.exp"
```
Chapter 7: Testsuites

A test that fails when the source files are compiled with different compilers, but passes when the files are compiled with the same compiler, demonstrates incompatibility of the generated code or runtime support. A test that fails for the alternate compiler but passes for the compiler under test probably tests for a bug that was fixed in the compiler under test but is present in the alternate compiler.

The binary compatibility tests support a small number of test framework commands that appear within comments in a test file.

dg-require-*
These commands can be used in 'testname_main.suffix' to skip the test if specific support is not available on the target.

dg-options
The specified options are used for compiling this particular source file, appended to the options from COMPAT_OPTIONS. When this command appears in 'testname_main.suffix' the options are also used to link the test program.

dg-xfail-if
This command can be used in a secondary source file to specify that compilation is expected to fail for particular options on particular targets.

7.10 Support for torture testing using multiple options

Throughout the compiler testsuite there are several directories whose tests are run multiple times, each with a different set of options. These are known as torture tests. ‘lib/torture-options.exp’ defines procedures to set up these lists:

```
torture-init
    Initialize use of torture lists.

set-torture-options
    Set lists of torture options to use for tests with and without loops. Optionally combine a set of torture options with a set of other options, as is done with Objective-C runtime options.

set-torture-finish
    Finalize use of torture lists.
```

The ‘.exp’ file for a set of tests that use torture options must include calls to these three procedures if:

- It calls gdc-dg-runtest and overrides DG_TORTURE_OPTIONS.
- It calls ${tool}-torture or ${tool}-torture-execute, where tool is c, fortran, or objc.
- It calls dg-pch.

It is not necessary for a ‘.exp’ file that calls gdc-dg-runtest to call the torture procedures if the tests should use the list in DG_TORTURE_OPTIONS defined in ‘gdc-dg.exp’.

Most uses of torture options can override the default lists by defining TORTURE_OPTIONS or add to the default list by defining ADDITIONAL_TORTURE_OPTIONS. Define these in a ‘.dejagnurc’ file or add them to the ‘site.exp’ file; for example
set ADDITIONAL_TORTURE_OPTIONS [list \\
{ -O2 -ftree-loop-linear } \\
{ -O2 -fpeel-loops } ]
8 Option specification files

Most GCC command-line options are described by special option definition files, the names of which conventionally end in `.opt`. This chapter describes the format of these files.

8.1 Option file format

Option files are a simple list of records in which each field occupies its own line and in which the records themselves are separated by blank lines. Comments may appear on their own line anywhere within the file and are preceded by semicolons. Whitespace is allowed before the semicolon.

The files can contain the following types of record:

- A language definition record. These records have two fields: the string `Language` and the name of the language. Once a language has been declared in this way, it can be used as an option property. See Section 8.2 [Option properties], page 105.

- A target specific save record to save additional information. These records have two fields: the string `TargetSave`, and a declaration type to go in the `cl_target_option` structure.

- A variable record to define a variable used to store option information. These records have two fields: the string `Variable`, and a declaration of the type and name of the variable, optionally with an initializer (but without any trailing `;`). These records may be used for variables used for many options where declaring the initializer in a single option definition record, or duplicating it in many records, would be inappropriate, or for variables set in option handlers rather than referenced by `Var` properties.

- A variable record to define a variable used to store option information. These records have two fields: the string `TargetVariable`, and a declaration of the type and name of the variable, optionally with an initializer (but without any trailing `;`). `TargetVariable` is a combination of `Variable` and `TargetSave` records in that the variable is defined in the `gcc_options` structure, but these variables are also stored in the `cl_target_option` structure. The variables are saved in the target save code and restored in the target restore code.

- A variable record to record any additional files that the `options.h` file should include. This is useful to provide enumeration or structure definitions needed for target variables. These records have two fields: the string `HeaderInclude` and the name of the include file.

- A variable record to record any additional files that the `options.c` file should include. This is useful to provide inline functions needed for target variables and/or `#ifdef` sequences to properly set up the initialization. These records have two fields: the string `SourceInclude` and the name of the include file.

- An enumeration record to define a set of strings that may be used as arguments to an option or options. These records have three fields: the string `Enum`, a space-separated list of properties and help text used to describe the set of strings in `--help` output. Properties use the same format as option properties; the following are valid:

  \[Name(name)\]

  This property is required; `name` must be a name (suitable for use in C identifiers) used to identify the set of strings in `Enum` option properties.
Type(type)
This property is required; type is the C type for variables set by options using this enumeration together with Var.

UnknownError(message)
The message message will be used as an error message if the argument is invalid; for enumerations without UnknownError, a generic error message is used. message should contain a single ‘%qs’ format, which will be used to format the invalid argument.

• An enumeration value record to define one of the strings in a set given in an ‘Enum’ record. These records have two fields: the string ‘EnumValue’ and a space-separated list of properties. Properties use the same format as option properties; the following are valid:

Enum(name)
This property is required; name says which ‘Enum’ record this ‘EnumValue’ record corresponds to.

String(string)
This property is required; string is the string option argument being described by this record.

Value(value)
This property is required; it says what value (representable as int) should be used for the given string.

Canonical
This property is optional. If present, it says the present string is the canonical one among all those with the given value. Other strings yielding that value will be mapped to this one so specs do not need to handle them.

DriverOnly
This property is optional. If present, the present string will only be accepted by the driver. This is used for cases such as ‘-march=native’ that are processed by the driver so that ‘gcc -v’ shows how the options chosen depended on the system on which the compiler was run.

• An option definition record. These records have the following fields:
  1. the name of the option, with the leading “-” removed
  2. a space-separated list of option properties (see Section 8.2 [Option properties], page 105)
  3. the help text to use for ‘--help’ (omitted if the second field contains the Undocumented property).

By default, all options beginning with “f”, “W” or “m” are implicitly assumed to take a “no-” form. This form should not be listed separately. If an option beginning with one of these letters does not have a “no-” form, you can use the RejectNegative property to reject it.

The help text is automatically line-wrapped before being displayed. Normally the name of the option is printed on the left-hand side of the output and the help text is printed
on the right. However, if the help text contains a tab character, the text to the left of the tab is used instead of the option’s name and the text to the right of the tab forms the help text. This allows you to elaborate on what type of argument the option takes.

- A target mask record. These records have one field of the form ‘\(\text{Mask}(x)\)’. The options-processing script will automatically allocate a bit in \(\text{target\_flags}\) (see Section 17.3 [Run-time Target], page 400) for each mask name \(x\) and set the macro \(\text{MASK}_x\) to the appropriate bitmask. It will also declare a \(\text{TARGET}_x\) macro that has the value 1 when bit \(\text{MASK}_x\) is set and 0 otherwise.

They are primarily intended to declare target masks that are not associated with user options, either because these masks represent internal switches or because the options are not available on all configurations and yet the masks always need to be defined.

8.2 Option properties

The second field of an option record can specify any of the following properties. When an option takes an argument, it is enclosed in parentheses following the option property name. The parser that handles option files is quite simplistic, and will be tricked by any nested parentheses within the argument text itself; in this case, the entire option argument can be wrapped in curly braces within the parentheses to demarcate it, e.g.:

\[
\text{Condition}\{\text{defined (USE_CYGWIN_LIBSTDCXX.WRAPPERS)}\}
\]

Common The option is available for all languages and targets.

Target The option is available for all languages but is target-specific.

Driver The option is handled by the compiler driver using code not shared with the compilers proper (‘\(\text{cc1}\)’ etc.).

\textbf{language} The option is available when compiling for the given language.

It is possible to specify several different languages for the same option. Each language must have been declared by an earlier Language record. See Section 8.1 [Option file format], page 103.

RejectDriver The option is only handled by the compilers proper (‘\(\text{cc1}\)’ etc.) and should not be accepted by the driver.

RejectNegative The option does not have a “no-” form. All options beginning with “\(f\)”, “\(W\)” or “\(m\)” are assumed to have a “no-” form unless this property is used.

\textbf{Negative(othername)} The option will turn off another option othername, which is the option name with the leading “\(-\)” removed. This chain action will propagate through the Negative property of the option to be turned off.

Joined The option takes a mandatory argument. Joined indicates that the option and argument can be included in the same argv entry (as with -mflush-func=name, for example). Separate indicates that the option and argument can be separate argv entries (as with -o). An option is allowed to have both of these properties.
**JoinedOrMissing**

The option takes an optional argument. If the argument is given, it will be part of the same `argv` entry as the option itself.

This property cannot be used alongside `Joined` or `Separate`.

**MissingArgError** *(message)*

For an option marked `Joined` or `Separate`, the message `message` will be used as an error message if the mandatory argument is missing; for options without `MissingArgError`, a generic error message is used. `message` should contain a single `\%qs` format, which will be used to format the name of the option passed.

**Args** *(n)*

For an option marked `Separate`, indicate that it takes `n` arguments. The default is 1.

**UInteger**

The option’s argument is a non-negative integer. The option parser will check and convert the argument before passing it to the relevant option handler. `UInteger` should also be used on options like `-falign-loops` where both `-falign-loops` and `-falign-loops=n` are supported to make sure the saved options are given a full integer.

**NoDriverArg**

For an option marked `Separate`, the option only takes an argument in the compiler proper, not in the driver. This is for compatibility with existing options that are used both directly and via `-Wp,`; new options should not have this property.

**Var** *(var)*

The state of this option should be stored in variable `var` (actually a macro for `global_options.x_var`). The way that the state is stored depends on the type of option:

- If the option uses the `Mask` or `InverseMask` properties, `var` is the integer variable that contains the mask.
- If the option is a normal on/off switch, `var` is an integer variable that is nonzero when the option is enabled. The options parser will set the variable to 1 when the positive form of the option is used and 0 when the “no-” form is used.
- If the option takes an argument and has the `UInteger` property, `var` is an integer variable that stores the value of the argument.
- If the option takes an argument and has the `Enum` property, `var` is a variable (type given in the `Type` property of the ‘Enum’ record whose `Name` property has the same argument as the `Enum` property of this option) that stores the value of the argument.
- If the option has the `Defer` property, `var` is a pointer to a `VEC(cl_deferred_option,heap)` that stores the option for later processing. (`var` is declared with type `void *` and needs to be cast to `VEC(cl_deferred_option,heap)` before use.)
- Otherwise, if the option takes an argument, `var` is a pointer to the argument string. The pointer will be null if the argument is optional and wasn’t given.
The option-processing script will usually zero-initialize var. You can modify this behavior using Init.

**Var(var, set)**

The option controls an integer variable var and is active when var equals set. The option parser will set var to set when the positive form of the option is used and !set when the “no-” form is used.

var is declared in the same way as for the single-argument form described above.

**Init(value)**

The variable specified by the Var property should be statically initialized to value. If more than one option using the same variable specifies Init, all must specify the same initializer.

**Mask(name)**

The option is associated with a bit in the target_flags variable (see Section 17.3 [Run-time Target], page 400) and is active when that bit is set. You may also specify Var to select a variable other than target_flags.

The options-processing script will automatically allocate a unique bit for the option. If the option is attached to 'target_flags', the script will set the macro MASK_name to the appropriate bitmask. It will also declare a TARGET_name macro that has the value 1 when the option is active and 0 otherwise. If you use Var to attach the option to a different variable, the associated macros are called OPTION_MASK_name and OPTION_name respectively.

You can disable automatic bit allocation using MaskExists.

**InverseMask(othername)**

**InverseMask(othername, thisname)**

The option is the inverse of another option that has the Mask(othername) property. If thisname is given, the options-processing script will declare a TARGET_thisname macro that is 1 when the option is active and 0 otherwise.

**MaskExists**

The mask specified by the Mask property already exists. No MASK or TARGET definitions should be added to 'options.h' in response to this option record.

The main purpose of this property is to support synonymous options. The first option should use 'Mask(othername)' and the others should use 'Mask(name) MaskExists'.

**Enum(name)**

The option’s argument is a string from the set of strings associated with the corresponding ‘Enum’ record. The string is checked and converted to the integer specified in the corresponding ‘EnumValue’ record before being passed to option handlers.

**Defer**

The option should be stored in a vector, specified with Var, for later processing.
**Aliases**

- **Alias(opt)**
  - The option is an alias for `-opt`. In the first form, any argument passed to the alias is considered to be passed to `-opt`, and `-opt` is considered to be negated if the alias is used in negated form. In the second form, the alias may not be negated or have an argument, and `posarg` is considered to be passed as an argument to `-opt`. In the third form, the alias may not have an argument, if the alias is used in the positive form then `posarg` is considered to be passed to `-opt`, and if the alias is used in the negative form then `negarg` is considered to be passed to `-opt`.

Aliases should not specify `Var` or `Mask` or `UInteger`. Aliases should normally specify the same languages as the target of the alias; the flags on the target will be used to determine any diagnostic for use of an option for the wrong language, while those on the alias will be used to identify what command-line text is the option and what text is any argument to that option.

When an **Alias** definition is used for an option, driver specs do not need to handle it and no ‘OPT_’ enumeration value is defined for it; only the canonical form of the option will be seen in those places.

- **Ignore**
  - This option is ignored apart from printing any warning specified using **Warn**. The option will not be seen by specs and no ‘OPT_’ enumeration value is defined for it.

- **SeparateAlias**
  - For an option marked with **Joined**, **Separate** and **Alias**, the option only acts as an alias when passed a separate argument; with a joined argument it acts as a normal option, with an ‘OPT_’ enumeration value. This is for compatibility with the Java ‘-d’ option and should not be used for new options.

- **Warn(message)**
  - If this option is used, output the warning `message`. `message` is a format string, either taking a single operand with a ‘%qs’ format which is the option name, or not taking any operands, which is passed to the ‘warning’ function. If an alias is marked **Warn**, the target of the alias must not also be marked **Warn**.

- **Report**
  - The state of the option should be printed by ‘-fverbose-asm’.

- **Warning**
  - This is a warning option and should be shown as such in ‘--help’ output. This flag does not currently affect anything other than ‘--help’.

- **Optimization**
  - This is an optimization option. It should be shown as such in ‘--help’ output, and any associated variable named using `Var` should be saved and restored when the optimization level is changed with `optimize` attributes.

- **Undocumented**
  - The option is deliberately missing documentation and should not be included in the ‘--help’ output.
Condition(\texttt{cond})

The option should only be accepted if preprocessor condition \texttt{cond} is true. Note that any C declarations associated with the option will be present even if \texttt{cond} is false; \texttt{cond} simply controls whether the option is accepted and whether it is printed in the `--help` output.

Save

Build the \texttt{cl\_target\_option} structure to hold a copy of the option, add the functions \texttt{cl\_target\_option\_save} and \texttt{cl\_target\_option\_restore} to save and restore the options.

SetByCombined

The option may also be set by a combined option such as `\texttt{-ffast-math}`. This causes the \texttt{gcc\_options} struct to have a field \texttt{frontend\_set\_name}, where \texttt{name} is the name of the field holding the value of this option (without the leading \texttt{x}). This gives the front end a way to indicate that the value has been set explicitly and should not be changed by the combined option. For example, some front ends use this to prevent `\texttt{-ffast-math}` and `\texttt{-fno-fast-math}` from changing the value of `\texttt{-fmath-errno}` for languages that do not use \texttt{errno}. 
9 Passes and Files of the Compiler

This chapter is dedicated to giving an overview of the optimization and code generation passes of the compiler. In the process, it describes some of the language front end interface, though this description is no where near complete.

9.1 Parsing pass

The language front end is invoked only once, via `lang_hooks.parse_file`, to parse the entire input. The language front end may use any intermediate language representation deemed appropriate. The C front end uses GENERIC trees (see Chapter 11 [GENERIC], page 177), plus a double handful of language specific tree codes defined in `c-common.def`. The Fortran front end uses a completely different private representation.

At some point the front end must translate the representation used in the front end to a representation understood by the language-independent portions of the compiler. Current practice takes one of two forms. The C front end manually invokes the gimplifier (see Chapter 12 [GIMPLE], page 221) on each function, and uses the gimplifier callbacks to convert the language-specific tree nodes directly to GIMPLE before passing the function off to be compiled. The Fortran front end converts from a private representation to GENERIC, which is later lowered to GIMPLE when the function is compiled. Which route to choose probably depends on how well GENERIC (plus extensions) can be made to match up with the source language and necessary parsing data structures.

BUG: Gimplification must occur before nested function lowering, and nested function lowering must be done by the front end before passing the data off to cgraph.

TODO: Cgraph should control nested function lowering. It would only be invoked when it is certain that the outer-most function is used.

TODO: Cgraph needs a gimplify_function callback. It should be invoked when (1) it is certain that the function is used, (2) warning flags specified by the user require some amount of compilation in order to honor, (3) the language indicates that semantic analysis is not complete until gimplification occurs. Hum... this sounds overly complicated. Perhaps we should just have the front end gimplify always; in most cases it’s only one function call.

The front end needs to pass all function definitions and top level declarations off to the middle-end so that they can be compiled and emitted to the object file. For a simple procedural language, it is usually most convenient to do this as each top level declaration or definition is seen. There is also a distinction to be made between generating functional code and generating complete debug information. The only thing that is absolutely required for functional code is that function and data definitions be passed to the middle-end. For complete debug information, function, data and type declarations should all be passed as well.

In any case, the front end needs each complete top-level function or data declaration, and each data definition should be passed to `rest_of_decl_compilation`. Each complete type definition should be passed to `rest_of_type_compilation`. Each function definition should be passed to `cgraph_finalize_function`.

TODO: I know rest_of_compilation currently has all sorts of RTL generation semantics. I plan to move all code generation bits (both Tree and RTL) to compile_function. Should we hide cgraph from the front ends and move back to rest_of_compilation as the official
interface? Possibly we should rename all three interfaces such that the names match in some meaningful way and that is more descriptive than "rest_of".

The middle-end will, at its option, emit the function and data definitions immediately or queue them for later processing.

### 9.2 Gimplification pass

Gimplification is a whimsical term for the process of converting the intermediate representation of a function into the GIMPLE language (see Chapter 12 [GIMPLE], page 221). The term stuck, and so words like “gimplification”, “gimplify”, “gimplifier” and the like are sprinkled throughout this section of code.

While a front end may certainly choose to generate GIMPLE directly if it chooses, this can be a moderately complex process unless the intermediate language used by the front end is already fairly simple. Usually it is easier to generate GENERIC trees plus extensions and let the language-independent gimplifier do most of the work.

The main entry point to this pass is gimplify_function_tree located in 'gimplify.c'. From here we process the entire function gimplifying each statement in turn. The main workhorse for this pass is gimplify_expr. Approximately everything passes through here at least once, and it is from here that we invoke the lang_hooks.gimplify_expr callback.

The callback should examine the expression in question and return GS_UNHANDLED if the expression is not a language specific construct that requires attention. Otherwise it should alter the expression in some way to such that forward progress is made toward producing valid GIMPLE. If the callback is certain that the transformation is complete and the expression is valid GIMPLE, it should return GS_ALL_DONE. Otherwise it should return GS_OK, which will cause the expression to be processed again. If the callback encounters an error during the transformation (because the front end is relying on the gimplification process to finish semantic checks), it should return GS_ERROR.

### 9.3 Pass manager

The pass manager is located in ‘passes.c’, ‘tree-optimize.c’ and ‘tree-pass.h’. Its job is to run all of the individual passes in the correct order, and take care of standard bookkeeping that applies to every pass.

The theory of operation is that each pass defines a structure that represents everything we need to know about that pass—when it should be run, how it should be run, what intermediate language form or on-the-side data structures it needs. We register the pass to be run in some particular order, and the pass manager arranges for everything to happen in the correct order.

The actuality doesn’t completely live up to the theory at present. Command-line switches and timevar_id_t enumerations must still be defined elsewhere. The pass manager validates constraints but does not attempt to (re-)generate data structures or lower intermediate language form based on the requirements of the next pass. Nevertheless, what is present is useful, and a far sight better than nothing at all.

Each pass should have a unique name. Each pass may have its own dump file (for GCC debugging purposes). Passes with a name starting with a star do not dump anything. Sometimes passes are supposed to share a dump file / option name. To still give these
unique names, you can use a prefix that is delimited by a space from the part that is used
for the dump file / option name. E.g. When the pass name is "ud dce", the name used for
dump file/options is "dce".

TODO: describe the global variables set up by the pass manager, and a brief description
of how a new pass should use it. I need to look at what info RTL passes use first...

### 9.4 Tree SSA passes

The following briefly describes the Tree optimization passes that are run after gimplification
and what source files they are located in.

- **Remove useless statements**
  
  This pass is an extremely simple sweep across the gimple code in which we identify
  obviously dead code and remove it. Here we do things like simplify if statements
  with constant conditions, remove exception handling constructs surrounding code that
  obviously cannot throw, remove lexical bindings that contain no variables, and other
  assorted simplistic cleanups. The idea is to get rid of the obvious stuff quickly rather
  than wait until later when it’s more work to get rid of it. This pass is located in
  'tree-cfg.c' and described by `pass_remove_useless_stmts`.

- **Mudflap declaration registration**
  
  If mudflap (see Section “-fmudflap -fmudflapth -fmudflapir” in Using the GNU Compiler
  Collection (GCC)) is enabled, we generate code to register some variable declarations
  with the mudflap runtime. Specifically, the runtime tracks the lifetimes of those variable
  declarations that have their addresses taken, or whose bounds are unknown at compile
time (extern). This pass generates new exception handling constructs (try/finally),
and so must run before those are lowered. In addition, the pass enqueues declarations
of static variables whose lifetimes extend to the entire program. The pass is located in
'tree-mudflap.c' and is described by `pass_mudflap_1`.

- **OpenMP lowering**
  
  If OpenMP generation ('-fopenmp') is enabled, this pass lowers OpenMP constructs
  into GIMPLE.

  Lowering of OpenMP constructs involves creating replacement expressions for local
  variables that have been mapped using data sharing clauses, exposing the control flow
  of most synchronization directives and adding region markers to facilitate the creation
  of the control flow graph. The pass is located in 'omp-low.c' and is described by `pass_lower_omp`.

- **OpenMP expansion**
  
  If OpenMP generation ('-fopenmp') is enabled, this pass expands parallel regions
  into their own functions to be invoked by the thread library. The pass is located
  in 'omp-low.c' and is described by `pass_expand_omp`.

- **Lower control flow**
  
  This pass flattens if statements (COND_EXPR) and moves lexical bindings (BIND_EXPR)
out of line. After this pass, all if statements will have exactly two goto statements in
its then and else arms. Lexical binding information for each statement will be found
in TREE_BLOCK rather than being inferred from its position under a BIND_EXPR. This
pass is found in 'gimple-low.c' and is described by `pass_lower_cf`. 
Lower exception handling control flow
This pass decomposes high-level exception handling constructs (TRY_FINALLY_EXPR and TRY_CATCH_EXPR) into a form that explicitly represents the control flow involved. After this pass, lookup_stmt_eh_region will return a non-negative number for any statement that may have EH control flow semantics; examine tree_can_throw_internal or tree_can_throw_external for exact semantics. Exact control flow may be extracted from foreach_reachable_handler. The EH region nesting tree is defined in 'except.h' and built in 'except.c'. The lowering pass itself is in 'tree-eh.c' and is described by pass_lower_eh.

Build the control flow graph
This pass decomposes a function into basic blocks and creates all of the edges that connect them. It is located in 'tree-cfg.c' and is described by pass_build_cfg.

Find all referenced variables
This pass walks the entire function and collects an array of all variables referenced in the function, referenced_vars. The index at which a variable is found in the array is used as a UID for the variable within this function. This data is needed by the SSA rewriting routines. The pass is located in 'tree-dfa.c' and is described by pass_referenced_vars.

Enter static single assignment form
This pass rewrites the function such that it is in SSA form. After this pass, all is_gimple_reg variables will be referenced by SSA_NAME, and all occurrences of other variables will be annotated with VDEFS and VUSES; PHI nodes will have been inserted as necessary for each basic block. This pass is located in 'tree-ssa.c' and is described by pass_build_ssa.

Warn for uninitialized variables
This pass scans the function for uses of SSA_NAMEs that are fed by default definition. For non-parameter variables, such uses are uninitialized. The pass is run twice, before and after optimization (if turned on). In the first pass we only warn for uses that are positively uninitialized; in the second pass we warn for uses that are possibly uninitialized. The pass is located in 'tree-ssa.c' and is defined by pass_early_warn_uninitialized and pass_late_warn_uninitialized.

Dead code elimination
This pass scans the function for statements without side effects whose result is unused. It does not do memory life analysis, so any value that is stored in memory is considered used. The pass is run multiple times throughout the optimization process. It is located in 'tree-ssa-dce.c' and is described by pass_dce.

Dominator optimizations
This pass performs trivial dominator-based copy and constant propagation, expression simplification, and jump threading. It is run multiple times throughout the optimization process. It is located in 'tree-ssa-dom.c' and is described by pass_dominator.

Forward propagation of single-use variables
This pass attempts to remove redundant computation by substituting variables that are used once into the expression that uses them and seeing if the result can be simplified. It is located in 'tree-ssa-forwprop.c' and is described by pass_forwprop.
• Copy Renaming
  This pass attempts to change the name of compiler temporaries involved in copy operations such that SSA-normal can coalesce the copy away. When compiler temporaries are copies of user variables, it also renames the compiler temporary to the user variable resulting in better use of user symbols. It is located in ‘tree-ssa-copyrename.c’ and is described by pass_copyrename.

• PHI node optimizations
  This pass recognizes forms of PHI inputs that can be represented as conditional expressions and rewrites them into straight line code. It is located in ‘tree-ssa-phiopt.c’ and is described by pass_phiopt.

• May-alias optimization
  This pass performs a flow sensitive SSA-based points-to analysis. The resulting may-alias, must-alias, and escape analysis information is used to promote variables from in-memory addressable objects to non-aliased variables that can be renamed into SSA form. We also update the VDEF/VUSE memory tags for non-renameable aggregates so that we get fewer false kills. The pass is located in ‘tree-ssa-alias.c’ and is described by pass_may_alias.
  Interprocedural points-to information is located in ‘tree-ssa-structalias.c’ and described by pass_ipa_pta.

• Profiling
  This pass rewrites the function in order to collect runtime block and value profiling data. Such data may be fed back into the compiler on a subsequent run so as to allow optimization based on expected execution frequencies. The pass is located in ‘predict.c’ and is described by pass_profile.

• Lower complex arithmetic
  This pass rewrites complex arithmetic operations into their component scalar arithmetic operations. The pass is located in ‘tree-complex.c’ and is described by pass_lower_complex.

• Scalar replacement of aggregates
  This pass rewrites suitable non-aliased local aggregate variables into a set of scalar variables. The resulting scalar variables are rewritten into SSA form, which allows subsequent optimization passes to do a significantly better job with them. The pass is located in ‘tree-sra.c’ and is described by pass_sra.

• Dead store elimination
  This pass eliminates stores to memory that are subsequently overwritten by another store, without any intervening loads. The pass is located in ‘tree-ssa-dse.c’ and is described by pass_dse.

• Tail recursion elimination
  This pass transforms tail recursion into a loop. It is located in ‘tree-tailcall.c’ and is described by pass_tail_recursion.

• Forward store motion
  This pass sinks stores and assignments down the flowgraph closer to their use point. The pass is located in ‘tree-ssa-sink.c’ and is described by pass_sink_code.
• Partial redundancy elimination
  This pass eliminates partially redundant computations, as well as performing load motion. The pass is located in ‘tree-ssa-pre.c’ and is described by pass_pre. Just before partial redundancy elimination, if ‘-funsafe-math-optimizations’ is on, GCC tries to convert divisions to multiplications by the reciprocal. The pass is located in ‘tree-ssa-math-opts.c’ and is described by pass_cse_reciprocal.

• Full redundancy elimination
  This is a simpler form of PRE that only eliminates redundancies that occur an all paths. It is located in ‘tree-ssa-pre.c’ and described by pass_fre.

• Loop optimization
  The main driver of the pass is placed in ‘tree-ssa-loop.c’ and described by pass_loop.
  The optimizations performed by this pass are:
  Loop invariant motion. This pass moves only invariants that would be hard to handle on RTL level (function calls, operations that expand to nontrivial sequences of insns). With ‘-funswitch-loops’ it also moves operands of conditions that are invariant out of the loop, so that we can use just trivial invaraintness analysis in loop unswitching. The pass also includes store motion. The pass is implemented in ‘tree-ssa-loop-im.c’.
  Canonical induction variable creation. This pass creates a simple counter for number of iterations of the loop and replaces the exit condition of the loop using it, in case when a complicated analysis is necessary to determine the number of iterations. Later optimizations then may determine the number easily. The pass is implemented in ‘tree-ssa-loop-ivcanon.c’.
  Induction variable optimizations. This pass performs standard induction variable optimizations, including strength reduction, induction variable merging and induction variable elimination. The pass is implemented in ‘tree-ssa-loop-ivopts.c’.
  Loop unswitching. This pass moves the conditional jumps that are invariant out of the loops. To achieve this, a duplicate of the loop is created for each possible outcome of conditional jump(s). The pass is implemented in ‘tree-ssa-loop-unswitch.c’. This pass should eventually replace the RTL level loop unswitching in ‘loop-unswitch.c’, but currently the RTL level pass is not completely redundant yet due to deficiencies in tree level alias analysis.
  The optimizations also use various utility functions contained in ‘tree-ssa-loop-manip.c’, ‘cfgloop.c’, ‘cfgloopanal.c’ and ‘cfgloopmanip.c’.
  Vectorization. This pass transforms loops to operate on vector types instead of scalar types. Data parallelism across loop iterations is exploited to group data elements from consecutive iterations into a vector and operate on them in parallel. Depending on available target support the loop is conceptually unrolled by a factor VF (vectorization factor), which is the number of elements operated upon in parallel in each iteration, and the VF copies of each scalar operation are fused to form a vector operation. Additional loop transformations such as peeling and versioning may take place to align the number of iterations, and to align the memory accesses in the loop. The pass is implemented in ‘tree-vectorizer.c’ (the main driver), ‘tree-vect-loop.c’ and ‘tree-vect-loop-manip.c’ (loop specific parts and general

SLP Vectorization. This pass performs vectorization of straight-line code. The pass is implemented in ‘tree-vectorizer.c’ (the main driver), ‘tree-vect-slp.c’, ‘tree-vect-stmts.c’ and ‘tree-vect-data-refs.c’.

Autoparallelization. This pass splits the loop iteration space to run into several threads. The pass is implemented in ‘tree-parloops.c’.

Graphite is a loop transformation framework based on the polyhedral model. Graphite stands for Gimple Represented as Polyhedra. The internals of this infrastructure are documented in http://gcc.gnu.org/wiki/Graphite. The passes working on this representation are implemented in the various ‘graphite-*’ files.

- Tree level if-conversion for vectorizer
  This pass applies if-conversion to simple loops to help vectorizer. We identify if convertible loops, if-convert statements and merge basic blocks in one big block. The idea is to present loop in such form so that vectorizer can have one to one mapping between statements and available vector operations. This pass is located in ‘tree-if-conv.c’ and is described by pass_if_conversion.

- Conditional constant propagation
  This pass relaxes a lattice of values in order to identify those that must be constant even in the presence of conditional branches. The pass is located in ‘tree-ssa-ccp.c’ and is described by pass_ccp.

  A related pass that works on memory loads and stores, and not just register values, is located in ‘tree-ssa-ccp.c’ and described by pass_store_ccp.

- Conditional copy propagation
  This is similar to constant propagation but the lattice of values is the “copy-of” relation. It eliminates redundant copies from the code. The pass is located in ‘tree-ssa-copy.c’ and described by pass_copy_prop.

  A related pass that works on memory copies, and not just register copies, is located in ‘tree-ssa-copy.c’ and described by pass_store_copy_prop.

- Value range propagation
  This transformation is similar to constant propagation but instead of propagating single constant values, it propagates known value ranges. The implementation is based on Patterson’s range propagation algorithm (Accurate Static Branch Prediction by Value Range Propagation, J. R. C. Patterson, PLDI ’95). In contrast to Patterson’s algorithm, this implementation does not propagate branch probabilities nor it uses more than a single range per SSA name. This means that the current implementation cannot be used for branch prediction (though adapting it would not be difficult). The pass is located in ‘tree-vrp.c’ and is described by pass_vrp.

- Folding built-in functions
  This pass simplifies built-in functions, as applicable, with constant arguments or with inferable string lengths. It is located in ‘tree-ssa-ccp.c’ and is described by pass_fold_builtins.
• Split critical edges

This pass identifies critical edges and inserts empty basic blocks such that the edge is no longer critical. The pass is located in `tree-cfg.c` and is described by `pass_split_crit_edges`.

• Control dependence dead code elimination

This pass is a stronger form of dead code elimination that can eliminate unnecessary control flow statements. It is located in `tree-ssa-dce.c` and is described by `pass_cd_dce`.

• Tail call elimination

This pass identifies function calls that may be rewritten into jumps. No code transformation is actually applied here, but the data and control flow problem is solved. The code transformation requires target support, and so is delayed until RTL. In the meantime `CALL_EXPR_TAILCALL` is set indicating the possibility. The pass is located in `tree-tailcall.c` and is described by `pass_tail_calls`. The RTL transformation is handled by `fixup_tail_calls` in `calls.c`.

• Warn for function return without value

For non-void functions, this pass locates return statements that do not specify a value and issues a warning. Such a statement may have been injected by falling off the end of the function. This pass is run last so that we have as much time as possible to prove that the statement is not reachable. It is located in `tree-cfg.c` and is described by `pass_warn_function_return`.

• Mudflap statement annotation

If mudflap is enabled, we rewrite some memory accesses with code to validate that the memory access is correct. In particular, expressions involving pointer dereferences (INDIRECT_REF, ARRAY_REF, etc.) are replaced by code that checks the selected address range against the mudflap runtime’s database of valid regions. This check includes an inline lookup into a direct-mapped cache, based on shift/mask operations of the pointer value, with a fallback function call into the runtime. The pass is located in `tree-mudflap.c` and is described by `pass_mudflap_2`.

• Leave static single assignment form

This pass rewrites the function such that it is in normal form. At the same time, we eliminate as many single-use temporaries as possible, so the intermediate language is no longer GIMPLE, but GENERIC. The pass is located in `tree-outof-ssa.c` and is described by `pass_del_ssa`.

• Merge PHI nodes that feed into one another

This is part of the CFG cleanup passes. It attempts to join PHI nodes from a forwarder CFG block into another block with PHI nodes. The pass is located in `tree-cfgcleanup.c` and is described by `pass_merge_phi`.

• Return value optimization

If a function always returns the same local variable, and that local variable is an aggregate type, then the variable is replaced with the return value for the function (i.e., the function’s DECL_RESULT). This is equivalent to the C++ named return value optimization applied to GIMPLE. The pass is located in `tree-nrv.c` and is described by `pass_nrv`. 
• Return slot optimization
If a function returns a memory object and is called as \texttt{var = foo()}, this pass tries to change the call so that the address of \texttt{var} is sent to the caller to avoid an extra memory copy. This pass is located in \texttt{tree-nrv.c} and is described by \texttt{pass\_return\_slot}.

• Optimize calls to \texttt{__builtin\_object\_size}
This is a propagation pass similar to CCP that tries to remove calls to \texttt{__builtin\_object\_size} when the size of the object can be computed at compile-time. This pass is located in \texttt{tree-object-size.c} and is described by \texttt{pass\_object\_sizes}.

• Loop invariant motion
This pass removes expensive loop-invariant computations out of loops. The pass is located in \texttt{tree-ssa-loop.c} and described by \texttt{pass\_lim}.

• Loop nest optimizations
This is a family of loop transformations that works on loop nests. It includes loop interchange, scaling, skewing and reversal and they are all geared to the optimization of data locality in array traversals and the removal of dependencies that hamper optimizations such as loop parallelization and vectorization. The pass is located in \texttt{tree-loop-linear.c} and described by \texttt{pass\_linear\_transform}.

• Removal of empty loops
This pass removes loops with no code in them. The pass is located in \texttt{tree-ssa-loop-ivcanon.c} and described by \texttt{pass\_empty\_loop}.

• Unrolling of small loops
This pass completely unrolls loops with few iterations. The pass is located in \texttt{tree-ssa-loop-ivcanon.c} and described by \texttt{pass\_complete\_unroll}.

• Predictive commoning
This pass makes the code reuse the computations from the previous iterations of the loops, especially loads and stores to memory. It does so by storing the values of these computations to a bank of temporary variables that are rotated at the end of loop. To avoid the need for this rotation, the loop is then unrolled and the copies of the loop body are rewritten to use the appropriate version of the temporary variable. This pass is located in \texttt{tree-predcom.c} and described by \texttt{pass\_predcom}.

• Array prefetching
This pass issues prefetch instructions for array references inside loops. The pass is located in \texttt{tree-ssa-loop-prefetch.c} and described by \texttt{pass\_loop\_prefetch}.

• Reassociation
This pass rewrites arithmetic expressions to enable optimizations that operate on them, like redundancy elimination and vectorization. The pass is located in \texttt{tree-ssa-reassoc.c} and described by \texttt{pass\_reassoc}.

• Optimization of \texttt{stdarg} functions
This pass tries to avoid the saving of register arguments into the stack on entry to \texttt{stdarg} functions. If the function doesn’t use any \texttt{va\_start} macros, no registers need to be saved. If \texttt{va\_start} macros are used, the \texttt{va\_list} variables don’t escape the function, it is only necessary to save registers that will be used in \texttt{va\_arg} macros. For instance, if \texttt{va\_arg} is only used with integral types in the function, floating point
registers don’t need to be saved. This pass is located in `tree-stdarg.c` and described by `pass_stdarg`.

9.5 RTL passes

The following briefly describes the RTL generation and optimization passes that are run after the Tree optimization passes.

- RTL generation

  The source files for RTL generation include `stmt.c`, `calls.c`, `expr.c`, `explow.c`, `expmed.c`, `function.c`, `optabs.c` and `emit-rtl.c`. Also, the file `insn-emit.c`, generated from the machine description by the program `genemit`, is used in this pass. The header file `expr.h` is used for communication within this pass.

  The header files `insn-flags.h` and `insn-codes.h`, generated from the machine description by the programs `genflags` and `gencodes`, tell this pass which standard names are available for use and which patterns correspond to them.

- Generation of exception landing pads

  This pass generates the glue that handles communication between the exception handling library routines and the exception handlers within the function. Entry points in the function that are invoked by the exception handling library are called landing pads. The code for this pass is located in `except.c`.

- Control flow graph cleanup

  This pass removes unreachable code, simplifies jumps to next, jumps to jump, jumps across jumps, etc. The pass is run multiple times. For historical reasons, it is occasionally referred to as the “jump optimization pass”. The bulk of the code for this pass is in `cfgcleanup.c`, and there are support routines in `cfgrtl.c` and `jump.c`.

- Forward propagation of single-def values

  This pass attempts to remove redundant computation by substituting variables that come from a single definition, and seeing if the result can be simplified. It performs copy propagation and addressing mode selection. The pass is run twice, with values being propagated into loops only on the second run. The code is located in `fwprop.c`.

- Common subexpression elimination

  This pass removes redundant computation within basic blocks, and optimizes addressing modes based on cost. The pass is run twice. The code for this pass is located in `cse.c`.

- Global common subexpression elimination

  This pass performs two different types of GCSE depending on whether you are optimizing for size or not (LCM based GCSE tends to increase code size for a gain in speed, while Morel-Renvoise based GCSE does not). When optimizing for size, GCSE is done using Morel-Renvoise Partial Redundancy Elimination, with the exception that it does not try to move invariants out of loops—that is left to the loop optimization pass. If MR PRE GCSE is done, code hoisting (aka unification) is also done, as well as load motion. If you are optimizing for speed, LCM (lazy code motion) based GCSE is done. LCM is based on the work of Knoop, Ruthing, and Steffen. LCM based GCSE also does loop invariant code motion. We also perform load and store motion when optimizing
for speed. Regardless of which type of GCSE is used, the GCSE pass also performs global constant and copy propagation. The source file for this pass is `gcse.c`, and the LCM routines are in `lcm.c`.

- **Loop optimization**
  This pass performs several loop related optimizations. The source files `cfgloopanal.c` and `cfgloopmanip.c` contain generic loop analysis and manipulation code. Initialization and finalization of loop structures is handled by `loop-init.c`. A loop invariant motion pass is implemented in `loop-invariant.c`. Basic block level optimizations—unrolling, peeling and unswitching loops—are implemented in `loop-unswitch.c` and `loop-unroll.c`. Replacing of the exit condition of loops by special machine-dependent instructions is handled by `loop-doloop.c`.

- **Jump bypassing**
  This pass is an aggressive form of GCSE that transforms the control flow graph of a function by propagating constants into conditional branch instructions. The source file for this pass is `gcse.c`.

- **If conversion**
  This pass attempts to replace conditional branches and surrounding assignments with arithmetic, boolean value producing comparison instructions, and conditional move instructions. In the very last invocation after reload, it will generate predicated instructions when supported by the target. The code is located in `ifcvt.c`.

- **Web construction**
  This pass splits independent uses of each pseudo-register. This can improve effect of the other transformation, such as CSE or register allocation. The code for this pass is located in `web.c`.

- **Instruction combination**
  This pass attempts to combine groups of two or three instructions that are related by data flow into single instructions. It combines the RTL expressions for the instructions by substitution, simplifies the result using algebra, and then attempts to match the result against the machine description. The code is located in `combine.c`.

- **Register movement**
  This pass looks for cases where matching constraints would force an instruction to need a reload, and this reload would be a register-to-register move. It then attempts to change the registers used by the instruction to avoid the move instruction. The code is located in `regmove.c`.

- **Mode switching optimization**
  This pass looks for instructions that require the processor to be in a specific “mode” and minimizes the number of mode changes required to satisfy all users. What these modes are, and what they apply to are completely target-specific. The code for this pass is located in `mode-switching.c`.

- **Modulo scheduling**
  This pass looks at innermost loops and reorders their instructions by overlapping different iterations. Modulo scheduling is performed immediately before instruction scheduling. The code for this pass is located in `modulo-sched.c`. 
• Instruction scheduling
This pass looks for instructions whose output will not be available by the time that it is used in subsequent instructions. Memory loads and floating point instructions often have this behavior on RISC machines. It re-orders instructions within a basic block to try to separate the definition and use of items that otherwise would cause pipeline stalls. This pass is performed twice, before and after register allocation. The code for this pass is located in ‘haifa-sched.c’, ‘sched-deps.c’, ‘sched-ebb.c’, ‘sched-rgn.c’ and ‘sched-vis.c’.

• Register allocation
These passes make sure that all occurrences of pseudo registers are eliminated, either by allocating them to a hard register, replacing them by an equivalent expression (e.g. a constant) or by placing them on the stack. This is done in several subpasses:
  • Register move optimizations. This pass makes some simple RTL code transformations which improve the subsequent register allocation. The source file is ‘regmove.c’.
  • The integrated register allocator (IRA). It is called integrated because coalescing, register live range splitting, and hard register preferencing are done on-the-fly during coloring. It also has better integration with the reload pass. Pseudo-registers spilled by the allocator or the reload have still a chance to get hard-registers if the reload evicts some pseudo-registers from hard-registers. The allocator helps to choose better pseudos for spilling based on their live ranges and to coalesce stack slots allocated for the spilled pseudo-registers. IRA is a regional register allocator which is transformed into Chaitin-Briggs allocator if there is one region. By default, IRA chooses regions using register pressure but the user can force it to use one region or regions corresponding to all loops.
    Source files of the allocator are ‘ira.c’, ‘ira-build.c’, ‘ira-costs.c’, ‘ira-conflicts.c’, ‘ira-color.c’, ‘ira-emit.c’, ‘ira-lives’, plus header files ‘ira.h’ and ‘ira-int.h’ used for the communication between the allocator and the rest of the compiler and between the IRA files.
  • Reloading. This pass renumbers pseudo registers with the hardware registers numbers they were allocated. Pseudo registers that did not get hard registers are replaced with stack slots. Then it finds instructions that are invalid because a value has failed to end up in a register, or has ended up in a register of the wrong kind. It fixes up these instructions by reloading the problematical values temporarily into registers. Additional instructions are generated to do the copying.
    The reload pass also optionally eliminates the frame pointer and inserts instructions to save and restore call-clobbered registers around calls.
    Source files are ‘reload.c’ and ‘reload1.c’, plus the header ‘reload.h’ used for communication between them.
  • Basic block reordering
This pass implements profile guided code positioning. If profile information is not available, various types of static analysis are performed to make the predictions normally coming from the profile feedback (IE execution frequency, branch probability, etc). It is implemented in the file ‘bb-reorder.c’, and the various prediction routines are in ‘predict.c’.
• Variable tracking
This pass computes where the variables are stored at each position in code and gen-
erates notes describing the variable locations to RTL code. The location lists are then
generated according to these notes to debug information if the debugging information
format supports location lists. The code is located in ‘var-tracking.c’.

• Delayed branch scheduling
This optional pass attempts to find instructions that can go into the delay slots of other
instructions, usually jumps and calls. The code for this pass is located in ‘reorg.c’.

• Branch shortening
On many RISC machines, branch instructions have a limited range. Thus, longer
sequences of instructions must be used for long branches. In this pass, the compiler
figures out what far each instruction will be from each other instruction, and
therefore whether the usual instructions, or the longer sequences, must be used for
each branch. The code for this pass is located in ‘final.c’.

• Register-to-stack conversion
Conversion from usage of some hard registers to usage of a register stack may be done
at this point. Currently, this is supported only for the floating-point registers of the
Intel 80387 coprocessor. The code for this pass is located in ‘reg-stack.c’.

• Final
This pass outputs the assembler code for the function. The source files are ‘final.c’
plus ‘insn-output.c’; the latter is generated automatically from the machine descrip-
tion by the tool ‘genoutput’. The header file ‘conditions.h’ is used for communication
between these files. If mudflap is enabled, the queue of deferred declarations and any
addressed constants (e.g., string literals) is processed by mudflap_finish_file into a
synthetic constructor function containing calls into the mudflap runtime.

• Debugging information output
This is run after final because it must output the stack slot offsets for pseudo registers
that did not get hard registers. Source files are ‘dbxout.c’ for DBX symbol table
format, ‘sdbout.c’ for SDB symbol table format, ‘dwarfout.c’ for DWARF symbol
table format, files ‘dwarf2out.c’ and ‘dwarf2asm.c’ for DWARF2 symbol table format,
and ‘vmsdbgout.c’ for VMS debug symbol table format.
Chapter 10: RTL Representation

10 RTL Representation

The last part of the compiler work is done on a low-level intermediate representation called Register Transfer Language. In this language, the instructions to be output are described, pretty much one by one, in an algebraic form that describes what the instruction does.

RTL is inspired by Lisp lists. It has both an internal form, made up of structures that point at other structures, and a textual form that is used in the machine description and in printed debugging dumps. The textual form uses nested parentheses to indicate the pointers in the internal form.

10.1 RTL Object Types

RTL uses five kinds of objects: expressions, integers, wide integers, strings and vectors. Expressions are the most important ones. An RTL expression (“RTX”, for short) is a C structure, but it is usually referred to with a pointer; a type that is given the typedef name rtx.

An integer is simply an int; their written form uses decimal digits. A wide integer is an integral object whose type is HOST_WIDE_INT; their written form uses decimal digits.

A string is a sequence of characters. In core it is represented as a char * in usual C fashion, and it is written in C syntax as well. However, strings in RTL may never be null. If you write an empty string in a machine description, it is represented in core as a null pointer rather than as a pointer to a null character. In certain contexts, these null pointers instead of strings are valid. Within RTL code, strings are most commonly found inside symbol_ref expressions, but they appear in other contexts in the RTL expressions that make up machine descriptions.

In a machine description, strings are normally written with double quotes, as you would in C. However, strings in machine descriptions may extend over many lines, which is invalid C, and adjacent string constants are not concatenated as they are in C. Any string constant may be surrounded with a single set of parentheses. Sometimes this makes the machine description easier to read.

There is also a special syntax for strings, which can be useful when C code is embedded in a machine description. Wherever a string can appear, it is also valid to write a C-style brace block. The entire brace block, including the outermost pair of braces, is considered to be the string constant. Double quote characters inside the braces are not special. Therefore, if you write string constants in the C code, you need not escape each quote character with a backslash.

A vector contains an arbitrary number of pointers to expressions. The number of elements in the vector is explicitly present in the vector. The written form of a vector consists of square brackets (‘[...]’) surrounding the elements, in sequence and with whitespace separating them. Vectors of length zero are not created; null pointers are used instead.

Expressions are classified by expression codes (also called RTX codes). The expression code is a name defined in ‘rtl.def’, which is also (in uppercase) a C enumeration constant. The possible expression codes and their meanings are machine-independent. The code of an RTX can be extracted with the macro GET_CODE (x) and altered with PUT_CODE (x, newcode).
The expression code determines how many operands the expression contains, and what kinds of objects they are. In RTL, unlike Lisp, you cannot tell by looking at an operand what kind of object it is. Instead, you must know from its context—from the expression code of the containing expression. For example, in an expression of code `subreg`, the first operand is to be regarded as an expression and the second operand as an integer. In an expression of code `plus`, there are two operands, both of which are to be regarded as expressions. In a `symbol_ref` expression, there is one operand, which is to be regarded as a string.

Expressions are written as parentheses containing the name of the expression type, its flags and machine mode if any, and then the operands of the expression (separated by spaces).

Expression code names in the ‘md’ file are written in lowercase, but when they appear in C code they are written in uppercase. In this manual, they are shown as follows: `const_int`.

In a few contexts a null pointer is valid where an expression is normally wanted. The written form of this is `(nil)`.

### 10.2 RTL Classes and Formats

The various expression codes are divided into several classes, which are represented by single characters. You can determine the class of an RTX code with the macro `GET_RTX_CLASS (code)`. Currently, ‘rtl.def’ defines these classes:

- **RTX_OBJ** An RTX code that represents an actual object, such as a register (REG) or a memory location (MEM, SYMBOL_REF). LO_SUM is also included; instead, SUBREG and STRICT_LOW_PART are not in this class, but in class x.

- **RTX_CONST_OBJ** An RTX code that represents a constant object. HIGH is also included in this class.

- **RTX_COMPARE** An RTX code for a non-symmetric comparison, such as GEU or LT.

- **RTX_COMM_COMPARE** An RTX code for a symmetric (commutative) comparison, such as EQ or ORDERED.

- **RTX_UNARY** An RTX code for a unary arithmetic operation, such as NEG, NOT, or ABS. This category also includes value extension (sign or zero) and conversions between integer and floating point.

- **RTX_COMM_ARITH** An RTX code for a commutative binary operation, such as PLUS or AND. NE and EQ are comparisons, so they have class <.

- **RTX_BIN_ARITH** An RTX code for a non-commutative binary operation, such as MINUS, DIV, or ASHIFTRT.
Chapter 10: RTL Representation

RTX_BITFIELD_OPS
An RTX code for a bit-field operation. Currently only ZERO_EXTRACT and SIGN_EXTRACT. These have three inputs and are lvalues (so they can be used for insertion as well). See Section 10.11 [Bit-Fields], page 156.

RTX_TERNARY
An RTX code for other three input operations. Currently only IF_THEN_ELSE, VEC_MERGE, SIGN_EXTRACT, ZERO_EXTRACT, and FMA.

RTXInsn
An RTX code for an entire instruction: INSN, JUMPInsn, and CALLInsn. See Section 10.19 [Insns], page 166.

RTX_MATCH
An RTX code for something that matches in insns, such as MATCH_DUP. These only occur in machine descriptions.

RTX_AUTOINC
An RTX code for an auto-increment addressing mode, such as POST_INC.

RTX_EXTRA
All other RTX codes. This category includes the remaining codes used only in machine descriptions (DEFINE_*, etc.). It also includes all the codes describing side effects (SET, USE, Clobber, etc.) and the non-insns that may appear on an insn chain, such as NOTE, BARRIER, and CODE_LABEL. SUBREG is also part of this class.

For each expression code, ‘rtl.def’ specifies the number of contained objects and their kinds using a sequence of characters called the format of the expression code. For example, the format of subreg is ‘ei’.

These are the most commonly used format characters:

e An expression (actually a pointer to an expression).
i An integer.
w A wide integer.
s A string.
E A vector of expressions.

A few other format characters are used occasionally:
u ‘u’ is equivalent to ‘e’ except that it is printed differently in debugging dumps. It is used for pointers to insns.
n ‘n’ is equivalent to ‘i’ except that it is printed differently in debugging dumps. It is used for the line number or code number of a note insn.
S ‘S’ indicates a string which is optional. In the RTL objects in core, ‘S’ is equivalent to ‘s’, but when the object is read, from an ‘md’ file, the string value of this operand may be omitted. An omitted string is taken to be the null string.
‘\( V \)’ indicates a vector which is optional. In the RTL objects in core, ‘\( V \)’ is equivalent to ‘\( E \)’, but when the object is read from an ‘\( md \)’ file, the vector value of this operand may be omitted. An omitted vector is effectively the same as a vector of no elements.

‘\( B \)’ indicates a pointer to basic block structure.

‘\( 0 \)’ means a slot whose contents do not fit any normal category. ‘\( 0 \)’ slots are not printed at all in dumps, and are often used in special ways by small parts of the compiler.

There are macros to get the number of operands and the format of an expression code:

\[
\text{GET_RTX_LENGTH (code)}
\]

Number of operands of an RTX of code \( \text{code} \).

\[
\text{GET_RTX_FORMAT (code)}
\]

The format of an RTX of code \( \text{code} \), as a C string.

Some classes of RTX codes always have the same format. For example, it is safe to assume that all comparison operations have format \( \text{ee} \).

\[
1 \quad \text{All codes of this class have format } \text{e}.
\]

\[
< \quad \text{All codes of these classes have format } \text{ee}.
\]

\[
c \quad \text{All codes of these classes have format } \text{e}\text{e}.
\]

\[
b \quad \text{All codes of these classes have format } \text{e}\text{e}\text{e}.
\]

\[
i \quad \text{All codes of this class have formats that begin with } \text{iueieie}. \text{ See Section 10.19 [Insns], page 166. Note that not all RTL objects linked onto an insn chain are of class } \text{i}.
\]

\[
o \quad \text{You can make no assumptions about the format of these codes.}
\]

10.3 Access to Operands

Operands of expressions are accessed using the macros \( \text{XEXP, XINT, XWINT and XSTR} \). Each of these macros takes two arguments: an expression-pointer (RTX) and an operand number (counting from zero). Thus,

\[
\text{XEXP (x, 2)}
\]

accesses operand 2 of expression \( x \), as an expression.

\[
\text{XINT (x, 2)}
\]

accesses the same operand as an integer. \( \text{XSTR} \), used in the same fashion, would access it as a string.

Any operand can be accessed as an integer, as an expression or as a string. You must choose the correct method of access for the kind of value actually stored in the operand. You would do this based on the expression code of the containing expression. That is also how you would know how many operands there are.
For example, if \( x \) is a \texttt{subreg} expression, you know that it has two operands which can be correctly accessed as \texttt{XEXP (x, 0)} and \texttt{XINT (x, 1)}. If you did \texttt{XINT (x, 0)}, you would get the address of the expression operand but cast as an integer; that might occasionally be useful, but it would be cleaner to write \((\texttt{int}) \texttt{XEXP (x, 0)}\). \texttt{XEXP (x, 1)} would also compile without error, and would return the second, integer operand cast as an expression pointer, which would probably result in a crash when accessed. Nothing stops you from writing \texttt{XEXP (x, 28)} either, but this will access memory past the end of the expression with unpredictable results.

Access to operands which are vectors is more complicated. You can use the macro \texttt{XVEC} to get the vector-pointer itself, or the macros \texttt{XVECEXP} and \texttt{XVECLEN} to access the elements and length of a vector.

\[
\texttt{XVEC (exp, idx)}
\]
Access the vector-pointer which is operand number \( idx \) in \( exp \).

\[
\texttt{XVECLEN (exp, idx)}
\]
Access the length (number of elements) in the vector which is in operand number \( idx \) in \( exp \). This value is an \texttt{int}.

\[
\texttt{XVECEXP (exp, idx, eltnum)}
\]
Access element number \( eltnum \) in the vector which is in operand number \( idx \) in \( exp \). This value is an \texttt{RTX}.

It is up to you to make sure that \( eltnum \) is not negative and is less than \texttt{XVECLEN (exp, idx)}.

All the macros defined in this section expand into lvalues and therefore can be used to assign the operands, lengths and vector elements as well as to access them.

### 10.4 Access to Special Operands

Some RTL nodes have special annotations associated with them.

\texttt{MEM}

\[
\texttt{MEM_ALIAS_SET (x)}
\]
If 0, \( x \) is not in any alias set, and may alias anything. Otherwise, \( x \) can only alias MEMs in a conflicting alias set. This value is set in a language-dependent manner in the front-end, and should not be altered in the back-end. In some front-ends, these numbers may correspond in some way to types, or other language-level entities, but they need not, and the back-end makes no such assumptions. These set numbers are tested with \texttt{alias_sets_conflict_p}.

\[
\texttt{MEM_EXPR (x)}
\]
If this register is known to hold the value of some user-level declaration, this is that tree node. It may also be a \texttt{COMPONENT_REF}, in which case this is some field reference, and \texttt{TREE_OPERAND (x, 0)} contains the declaration, or another \texttt{COMPONENT_REF}, or null if there is no compile-time object associated with the reference.

\[
\texttt{MEM_OFFSET (x)}
\]
The offset from the start of \texttt{MEM_EXPR} as a \texttt{CONST_INT} rtx.
MEM_SIZE (x)
The size in bytes of the memory reference as a CONST_INT rtx. This is mostly relevant for BLKmode references as otherwise the size is implied by the mode.

MEM_ALIGN (x)
The known alignment in bits of the memory reference.

MEM_ADDR_SPACE (x)
The address space of the memory reference. This will commonly be zero for the generic address space.

REG

ORIGINAL_REGNO (x)
This field holds the number the register “originally” had; for a pseudo register turned into a hard reg this will hold the old pseudo register number.

REG_EXPR (x)
If this register is known to hold the value of some user-level declaration, this is that tree node.

REG_OFFSET (x)
If this register is known to hold the value of some user-level declaration, this is the offset into that logical storage.

SYMBOL_REF

SYMBOL_REFDECL (x)
If the symbol_ref x was created for a VARDECL or a FUNCTION_DECL, that tree is recorded here. If this value is null, then x was created by back end code generation routines, and there is no associated front end symbol table entry.

SYMBOL_REFDECL may also point to a tree of class ‘c’, that is, some sort of constant. In this case, the symbol_ref is an entry in the per-file constant pool; again, there is no associated front end symbol table entry.

SYMBOLREFCONSTANT (x)
If ‘CONSTANT_POOL_ADDRESS_P (x)’ is true, this is the constant pool entry for x. It is null otherwise.

SYMBOL_REF_DATA (x)
A field of opaque type used to store SYMBOL_REFDECL or SYMBOL_REFCONSTANT.

SYMBOL_REF_FLAGS (x)
In a symbol_ref, this is used to communicate various predicates about the symbol. Some of these are common enough to be computed by common code, some are specific to the target. The common bits are:
SYMBOL_FLAG_FUNCTION
Set if the symbol refers to a function.

SYMBOL_FLAG_LOCAL
Set if the symbol is local to this “module”. See TARGET_BINDS_LOCAL_P.

SYMBOL_FLAG_EXTERNAL
Set if this symbol is not defined in this translation unit. Note that this is not the inverse of SYMBOL_FLAG_LOCAL.

SYMBOL_FLAG_SMALL
Set if the symbol is located in the small data section. See TARGET_IN_SMALL_DATA_P.

SYMBOL_FLAG_TLS_MODEL (x)
This is a multi-bit field accessor that returns the tls_model to be used for a thread-local storage symbol. It returns zero for non-thread-local symbols.

SYMBOL_FLAG_HAS_BLOCK_INFO
Set if the symbol has SYMBOL_REF_BLOCK and SYMBOL_REF_BLOCK_OFFSET fields.

SYMBOL_FLAG_ANCHOR
Set if the symbol is used as a section anchor. “Section anchors” are symbols that have a known position within an object_block and that can be used to access nearby members of that block. They are used to implement ‘-fsection-anchors’.
If this flag is set, then SYMBOL_FLAG_HAS_BLOCK_INFO will be too.

Bits beginning with SYMBOL_FLAG_MACH_DEP are available for the target’s use.

SYMBOL_REF_BLOCK (x)
If ‘SYMBOL_REF_HAS_BLOCK_INFO_P (x)’, this is the ‘object_block’ structure to which the symbol belongs, or NULL if it has not been assigned a block.

SYMBOL_REF_BLOCK_OFFSET (x)
If ‘SYMBOL_REF_HAS_BLOCK_INFO_P (x)’, this is the offset of x from the first object in ‘SYMBOL_REF_BLOCK (x)’. The value is negative if x has not yet been assigned to a block, or it has not been given an offset within that block.

10.5 Flags in an RTL Expression
RTL expressions contain several flags (one-bit bit-fields) that are used in certain types of expression. Most often they are accessed with the following macros, which expand into lvalues.
CONSTANT_POOL_ADDRESS_P (x)
Nonzero in a symbol_ref if it refers to part of the current function’s constant pool. For most targets these addresses are in a .rodata section entirely separate from the function, but for some targets the addresses are close to the beginning of the function. In either case GCC assumes these addresses can be addressed directly, perhaps with the help of base registers. Stored in the unchanging field and printed as ‘/u’.

RTL_CONST_CALL_P (x)
In a call_insn indicates that the insn represents a call to a const function. Stored in the unchanging field and printed as ‘/u’.

RTLPURE_CALL_P (x)
In a call_insn indicates that the insn represents a call to a pure function. Stored in the return_val field and printed as ‘/i’.

RTL_CONST_OR_PURE_CALL_P (x)
In a call_insn, true if RTL_CONST_CALL_P or RTL_PURE_CALL_P is true.

RTL_LOOPING_CONST_OR_PURE_CALL_P (x)
In a call_insn indicates that the insn represents a possibly infinite looping call to a const or pure function. Stored in the call field and printed as ‘/c’. Only true if one of RTL_CONST_CALL_P or RTL_PURE_CALL_P is true.

INSN_ANNUNLED_BRANCH_P (x)
In a jump_insn, call_insn, or insn indicates that the branch is an annulling one. See the discussion under sequence below. Stored in the unchanging field and printed as ‘/u’.

INSN_DELETED_P (x)
In an insn, call_insn, jump_insn, code_label, barrier, or note, nonzero if the insn has been deleted. Stored in the volatil field and printed as ‘/v’.

INSN_FROM_TARGET_P (x)
In an insn or jump_insn or call_insn in a delay slot of a branch, indicates that the insn is from the target of the branch. If the branch insn has INSN_ANNUNLED_BRANCH_P set, this insn will only be executed if the branch is taken. For annulled branches with INSN_FROM_TARGET_P clear, the insn will be executed only if the branch is not taken. When INSN_ANNUNLED_BRANCH_P is not set, this insn will always be executed. Stored in the in_struct field and printed as ‘/s’.

LABEL_PRESERVE_P (x)
In a code_label or note, indicates that the label is referenced by code or data not visible to the RTL of a given function. Labels referenced by a non-local goto will have this bit set. Stored in the in_struct field and printed as ‘/s’.

LABEL_REF_NONLOCAL_P (x)
In label_ref and reg_label expressions, nonzero if this is a reference to a non-local label. Stored in the volatil field and printed as ‘/v’.

MEM_IN_STRUCT_P (x)
In mem expressions, nonzero for reference to an entire structure, union or array, or to a component of one. Zero for references to a scalar variable or through
a pointer to a scalar. If both this flag and MEM_SCALAR_P are clear, then we
don’t know whether this mem is in a structure or not. Both flags should never
be simultaneously set. Stored in the in_struct field and printed as ‘/s’.

MEM_KEEP_ALIAS_SET_P (x)
In mem expressions, 1 if we should keep the alias set for this mem unchanged
when we access a component. Set to 1, for example, when we are already in
a non-addressable component of an aggregate. Stored in the jump field and
printed as ‘/j’.

MEM_SCALAR_P (x)
In mem expressions, nonzero for reference to a scalar known not to be a member
of a structure, union, or array. Zero for such references and for indirections
through pointers, even pointers pointing to scalar types. If both this flag and
MEM_IN_STRUCT_P are clear, then we don’t know whether this mem is in a struc-
ture or not. Both flags should never be simultaneously set. Stored in the
return_val field and printed as ‘/i’.

MEM_VOLATILE_P (x)
In mem, asm_operands, and asm_input expressions, nonzero for volatile memory
references. Stored in the volatile field and printed as ‘/v’.

MEM_NOTRAP_P (x)
In mem, nonzero for memory references that will not trap. Stored in the call
field and printed as ‘/c’.

MEM_POINTER (x)
Nonzero in a mem if the memory reference holds a pointer. Stored in the frame-
related field and printed as ‘/f’.

REG_FUNCTION_VALUE_P (x)
Nonzero in a reg if it is the place in which this function’s value is going to be
returned. (This happens only in a hard register.) Stored in the return_val
field and printed as ‘/i’.

REG_POINTER (x)
Nonzero in a reg if the register holds a pointer. Stored in the frame_related
field and printed as ‘/f’.

REG_USERVAR_P (x)
In a reg, nonzero if it corresponds to a variable present in the user’s source
code. Zero for temporaries generated internally by the compiler. Stored in the
volatile field and printed as ‘/v’.

The same hard register may be used also for collecting the values of functions
called by this one, but REG_FUNCTION_VALUE_P is zero in this kind of use.

RTX_FRAME_RELATED_P (x)
Nonzero in an insn, call_insn, jump_insn, barrier, or set which is part of a
function prologue and sets the stack pointer, sets the frame pointer, or saves a
register. This flag should also be set on an instruction that sets up a temporary
register to use in place of the frame pointer. Stored in the frame_related field
and printed as ‘/f’.
In particular, on RISC targets where there are limits on the sizes of immediate
constants, it is sometimes impossible to reach the register save area directly from
the stack pointer. In that case, a temporary register is used that is near enough
to the register save area, and the Canonical Frame Address, i.e., DWARF2’s
logical frame pointer, register must (temporarily) be changed to be this tem-
porary register. So, the instruction that sets this temporary register must be
marked as RTX_FRAME_RELATED_P.

If the marked instruction is overly complex (defined in terms of what
dwarf2out_frame_debug_expr can handle), you will also have to create a
REG_FRAME_RELATED_EXPR note and attach it to the instruction. This note
should contain a simple expression of the computation performed by this
instruction, i.e., one that dwarf2out_frame_debug_expr can handle.

This flag is required for exception handling support on targets with RTL pro-
logues.

MEM_READONLY_P (x)
Nonzero in a mem, if the memory is statically allocated and read-only.
Read-only in this context means never modified during the lifetime of the pro-
gram, not necessarily in ROM or in write-disabled pages. A common example
of the later is a shared library’s global offset table. This table is initialized by
the runtime loader, so the memory is technically writable, but after control is
transferred from the runtime loader to the application, this memory will never
be subsequently modified.

Stored in the unchanging field and printed as ‘/u’.

SCHED_GROUP_P (x)
During instruction scheduling, in an insn, call_insn or jump_insn, indicates
that the previous insn must be scheduled together with this insn. This is
used to ensure that certain groups of instructions will not be split up by the
instruction scheduling pass, for example, use insns before a call_insn may not
be separated from the call_insn. Stored in the in_struct field and printed
as ‘/s’.

SET_IS_RETURN_P (x)
For a set, nonzero if it is for a return. Stored in the jump field and printed as
’/j’.

SIBLING_CALL_P (x)
For a call_insn, nonzero if the insn is a sibling call. Stored in the jump field
and printed as ’/j’.

STRING_POOL_ADDRESS_P (x)
For a symbol_ref expression, nonzero if it addresses this function’s string con-
stant pool. Stored in the frame_related field and printed as ’/f’.

SUBREG_PROMOTED_UNSIGNED_P (x)
Returns a value greater then zero for a subreg that has SUBREG_PROMOTED_\nVAR_P nonzero if the object being referenced is kept zero-extended, zero if it
is kept sign-extended, and less then zero if it is extended some other way via
the ptr_extend instruction. Stored in the unchanging field and volatil field, printed as ‘/u’ and ‘/v’. This macro may only be used to get the value it may not be used to change the value. Use SUBREG_PROMOTED_UNSIGNED_SET to change the value.

SUBREG_PROMOTED_UNSIGNED_SET (x)
Set the unchanging and volatil fields in a subreg to reflect zero, sign, or other extension. If volatil is zero, then unchanging as nonzero means zero extension and as zero means sign extension. If volatil is nonzero then some other type of extension was done via the ptr_extend instruction.

SUBREG_PROMOTED_VAR_P (x)
Nonzero in a subreg if it was made when accessing an object that was promoted to a wider mode in accord with the PROMOTED_MODE machine description macro (see Section 17.5 [Storage Layout], page 404). In this case, the mode of the subreg is the declared mode of the object and the mode of SUBREG_REG is the mode of the register that holds the object. Promoted variables are always either sign- or zero-extended to the wider mode on every assignment. Stored in the in_struct field and printed as ‘/s’.

SYMBOL_REF_USED (x)
In a symbol_ref, indicates that x has been used. This is normally only used to ensure that x is only declared external once. Stored in the used field.

SYMBOL_REF_WEAK (x)
In a symbol_ref, indicates that x has been declared weak. Stored in the return_val field and printed as ‘/i’.

SYMBOL_REF_FLAG (x)
In a symbol_ref, this is used as a flag for machine-specific purposes. Stored in the volatil field and printed as ‘/v’.
Most uses of SYMBOL_REF_FLAG are historic and may be subsumed by SYMBOL_REF_FLAGS. Certainly use of SYMBOL_REF_FLAGS is mandatory if the target requires more than one bit of storage.

PREFETCH_SCHEDULE_BARRIER_P (x)
In a prefetch, indicates that the prefetch is a scheduling barrier. No other INSNs will be moved over it. Stored in the volatil field and printed as ‘/v’.

These are the fields to which the above macros refer:
call
In a mem, 1 means that the memory reference will not trap.
In a call, 1 means that this pure or const call may possibly infinite loop.
In an RTL dump, this flag is represented as ‘/c’.

frame_related
In an insn or set expression, 1 means that it is part of a function prologue and sets the stack pointer, sets the frame pointer, saves a register, or sets up a temporary register to use in place of the frame pointer.
In reg expressions, 1 means that the register holds a pointer.
In mem expressions, 1 means that the memory reference holds a pointer.
In `symbol_ref` expressions, 1 means that the reference addresses this function’s string constant pool.
In an RTL dump, this flag is represented as ‘/f’.

**in_struct**

In `mem` expressions, it is 1 if the memory datum referred to is all or part of a structure or array; 0 if it is (or might be) a scalar variable. A reference through a C pointer has 0 because the pointer might point to a scalar variable. This information allows the compiler to determine something about possible cases of aliasing.
In `reg` expressions, it is 1 if the register has its entire life contained within the test expression of some loop.
In `subreg` expressions, 1 means that the `subreg` is accessing an object that has had its mode promoted from a wider mode.
In `label_ref` expressions, 1 means that the referenced label is outside the innermost loop containing the insn in which the `label_ref` was found.
In `code_label` expressions, it is 1 if the label may never be deleted. This is used for labels which are the target of non-local gotos. Such a label that would have been deleted is replaced with a `note` of type `NOTE_INSN_DELETED_LABEL`.
In an `insn` during dead-code elimination, 1 means that the insn is dead code.
In an `insn` or `jump_insn` during reorg for an insn in the delay slot of a branch, 1 means that this insn is from the target of the branch.
In an `insn` during instruction scheduling, 1 means that this insn must be scheduled as part of a group together with the previous insn.
In an RTL dump, this flag is represented as ‘/u’.

**return_val**

In `reg` expressions, 1 means the register contains the value to be returned by the current function. On machines that pass parameters in registers, the same register number may be used for parameters as well, but this flag is not set on such uses.
In `mem` expressions, 1 means the memory reference is to a scalar known not to be a member of a structure, union, or array.
In `symbol_ref` expressions, 1 means the referenced symbol is weak.
In `call` expressions, 1 means the call is pure.
In an RTL dump, this flag is represented as ‘/i’.

**jump**

In a `mem` expression, 1 means we should keep the alias set for this mem unchanged when we access a component.
In a `set`, 1 means it is for a return.
In a `call_insn`, 1 means it is a sibling call.
In an RTL dump, this flag is represented as ‘/j’.

**unchanging**

In `reg` and `mem` expressions, 1 means that the value of the expression never changes.
In `subreg` expressions, it is 1 if the `subreg` references an unsigned object whose mode has been promoted to a wider mode.

In an `insn` or `jump_insn` in the delay slot of a branch instruction, 1 means an annulling branch should be used.

In a `symbol_ref` expression, 1 means that this symbol addresses something in the per-function constant pool.

In a `call_insn` 1 means that this instruction is a call to a const function.

In an RTL dump, this flag is represented as '/u'.

`used` This flag is used directly (without an access macro) at the end of RTL generation for a function, to count the number of times an expression appears in insns. Expressions that appear more than once are copied, according to the rules for shared structure (see Section 10.21 [Sharing], page 175).

For a `reg`, it is used directly (without an access macro) by the leaf register renumbering code to ensure that each register is only renumbered once.

In a `symbol_ref`, it indicates that an external declaration for the symbol has already been written.

`volatile` In a `mem`, `asm_operands`, or `asm_input` expression, it is 1 if the memory reference is volatile. Volatile memory references may not be deleted, reordered or combined.

In a `symbol_ref` expression, it is used for machine-specific purposes.

In a `reg` expression, it is 1 if the value is a user-level variable. 0 indicates an internal compiler temporary.

In an `insn`, 1 means the insn has been deleted.

In `label_ref` and `reg_label` expressions, 1 means a reference to a non-local label.

In `prefetch` expressions, 1 means that the containing insn is a scheduling barrier.

In an RTL dump, this flag is represented as '/v'.

### 10.6 Machine Modes

A machine mode describes a size of data object and the representation used for it. In the C code, machine modes are represented by an enumeration type, `enum machine_mode`, defined in `machmode.def`. Each RTL expression has room for a machine mode and so do certain kinds of tree expressions (declarations and types, to be precise).

In debugging dumps and machine descriptions, the machine mode of an RTL expression is written after the expression code with a colon to separate them. The letters 'mode' which appear at the end of each machine mode name are omitted. For example, `(reg:SI 38)` is a `reg` expression with machine mode `SImode`. If the mode is `VOIDmode`, it is not written at all.

Here is a table of machine modes. The term “byte” below refers to an object of `BITS_PER_UNIT` bits (see Section 17.5 [Storage Layout], page 404).

- **BImode**  "Bit" mode represents a single bit, for predicate registers.
QImode  “Quarter-Integer” mode represents a single byte treated as an integer.

HImode  “Half-Integer” mode represents a two-byte integer.

PSImode “Partial Single Integer” mode represents an integer which occupies four bytes but which doesn’t really use all four. On some machines, this is the right mode to use for pointers.

SImode  “Single Integer” mode represents a four-byte integer.

PDImode “Partial Double Integer” mode represents an integer which occupies eight bytes but which doesn’t really use all eight. On some machines, this is the right mode to use for certain pointers.

DImode  “Double Integer” mode represents an eight-byte integer.

TImode  “Tetra Integer” (?) mode represents a sixteen-byte integer.

OImode  “Octa Integer” (?) mode represents a thirty-two-byte integer.

QFmode “Quarter-Floating” mode represents a quarter-precision (single byte) floating point number.

HFmode “Half-Floating” mode represents a half-precision (two byte) floating point number.

TQFmode “Three-Quarter-Floating” (?) mode represents a three-quarter-precision (three byte) floating point number.

SFmode “Single Floating” mode represents a four byte floating point number. In the common case, of a processor with IEEE arithmetic and 8-bit bytes, this is a single-precision IEEE floating point number; it can also be used for double-precision (on processors with 16-bit bytes) and single-precision VAX and IBM types.

DFmode “Double Floating” mode represents an eight byte floating point number. In the common case, of a processor with IEEE arithmetic and 8-bit bytes, this is a double-precision IEEE floating point number.

XFmode “Extended Floating” mode represents an IEEE extended floating point number. This mode only has 80 meaningful bits (ten bytes). Some processors require such numbers to be padded to twelve bytes, others to sixteen; this mode is used for either.

SDmode “Single Decimal Floating” mode represents a four byte decimal floating point number (as distinct from conventional binary floating point).

DDmode “Double Decimal Floating” mode represents an eight byte decimal floating point number.

TDmode “Tetra Decimal Floating” mode represents a sixteen byte decimal floating point number all 128 of whose bits are meaningful.

TFmode “Tetra Floating” mode represents a sixteen byte floating point number all 128 of whose bits are meaningful. One common use is the IEEE quad-precision format.
QQmode  “Quarter-Fractional” mode represents a single byte treated as a signed fractional number. The default format is “s.7”.

HQmode  “Half-Fractional” mode represents a two-byte signed fractional number. The default format is “s.15”.

SQmode  “Single Fractional” mode represents a four-byte signed fractional number. The default format is “s.31”.

DQmode  “Double Fractional” mode represents an eight-byte signed fractional number. The default format is “s.63”.

TQmode  “Tetra Fractional” mode represents a sixteen-byte signed fractional number. The default format is “s.127”.

UQQmode “Unsigned Quarter-Fractional” mode represents a single byte treated as an unsigned fractional number. The default format is “.8”.

UHQmode “Unsigned Half-Fractional” mode represents a two-byte unsigned fractional number. The default format is “.16”.

USQmode “Unsigned Single Fractional” mode represents a four-byte unsigned fractional number. The default format is “.32”.

UDQmode “Unsigned Double Fractional” mode represents an eight-byte unsigned fractional number. The default format is “.64”.

UTQmode “Unsigned Tetra Fractional” mode represents a sixteen-byte unsigned fractional number. The default format is “.128”.

HAmode  “Half-Accumulator” mode represents a two-byte signed accumulator. The default format is “s8.7”.

SAmode  “Single Accumulator” mode represents a four-byte signed accumulator. The default format is “s16.15”.

DAmode  “Double Accumulator” mode represents an eight-byte signed accumulator. The default format is “s32.31”.

TAmode  “Tetra Accumulator” mode represents a sixteen-byte signed accumulator. The default format is “s64.63”.

UHAmode “Unsigned Half-Accumulator” mode represents a two-byte unsigned accumulator. The default format is “8.8”.

USAmode “Unsigned Single Accumulator” mode represents a four-byte unsigned accumulator. The default format is “16.16”.

UDAmode “Unsigned Double Accumulator” mode represents an eight-byte unsigned accumulator. The default format is “32.32”.

UTAmode “Unsigned Tetra Accumulator” mode represents a sixteen-byte unsigned accumulator. The default format is “64.64”.

CCmode  “Condition Code” mode represents the value of a condition code, which is a machine-specific set of bits used to represent the result of a comparison operation. Other machine-specific modes may also be used for the condition code.
These modes are not used on machines that use cc0 (see Section 17.16 [Condition Code], page 476).

**BLKmode**  
“Block” mode represents values that are aggregates to which none of the other modes apply. In RTL, only memory references can have this mode, and only if they appear in string-move or vector instructions. On machines which have no such instructions, BLKmode will not appear in RTL.

**VOIDmode**  
Void mode means the absence of a mode or an unspecified mode. For example, RTL expressions of code `const_int` have mode VOIDmode because they can be taken to have whatever mode the context requires. In debugging dumps of RTL, VOIDmode is expressed by the absence of any mode.

**QCmode, HCmode, SCmode, DCmode, XCmode, TCmode**  
These modes stand for a complex number represented as a pair of floating point values. The floating point values are in QFmode, HFmode, SFmode, DFmode, XFmode, and TFmode, respectively.

**CQImode, CHImode, CSImode, CDImode, CTImode, COImode**  
These modes stand for a complex number represented as a pair of integer values. The integer values are in QImode, HImode, SImode, DImode, TImode, and OImode, respectively.

The machine description defines `Pmode` as a C macro which expands into the machine mode used for addresses. Normally this is the mode whose size is `BITS_PER_WORD`, SImode on 32-bit machines.

The only modes which a machine description must support are QImode, and the modes corresponding to `BITS_PER_WORD`, FLOAT_TYPE_SIZE and DOUBLE_TYPE_SIZE. The compiler will attempt to use DImode for 8-byte structures and unions, but this can be prevented by overriding the definition of MAX_FIXED_MODE_SIZE. Alternatively, you can have the compiler use TImode for 16-byte structures and unions. Likewise, you can arrange for the C type `short int` to avoid using HImode.

Very few explicit references to machine modes remain in the compiler and these few references will soon be removed. Instead, the machine modes are divided into mode classes. These are represented by the enumeration type `enum mode_class` defined in ‘machmode.h’. The possible mode classes are:

**MODE_INT**  
Integer modes. By default these are BImode, QImode, HImode, SImode, DImode, TImode, and OImode.

**MODE_PARTIAL_INT**  
The “partial integer” modes, PQImode, PHImode, PSImode and PDImode.

**MODE_FLOAT**  
Floating point modes. By default these are QFmode, HFmode, TQFmode, SFmode, DFmode, XFmode and TFmode.

**MODE_DECIMAL_FLOAT**  
Decimal floating point modes. By default these are SDmode, DDmode and TDmode.

**MODE_FRACT**  
Signed fractional modes. By default these are QQmode, HQmode, SQmode, DQmode and TQmode.
MODE_UFRONT
Unsigned fractional modes. By default these are UQQmode, UHQmode, USQmode, UDQmode and UTQmode.

MODE_ACCUM
Signed accumulator modes. By default these are HAmode, SAmode, DAmode and TAmode.

MODE_UACCUM
Unsigned accumulator modes. By default these are UHAmode, USAmode, UDAmode and UTAmode.

MODE_COMPLEX_INT
Complex integer modes. (These are not currently implemented).

MODE_COMPLEX_FLOAT
Complex floating point modes. By default these are QCmode, HCmode, SCmode, DCmode, XCmode, and TCmode.

MODE_FUNCTION
Algol or Pascal function variables including a static chain. (These are not currently implemented).

MODE_CC
Modes representing condition code values. These are CCmode plus any CC_MODE modes listed in the `machine-modes.def`. See Section 16.12 [Jump Patterns], page 357, also see Section 17.16 [Condition Code], page 476.

MODE_RANDOM
This is a catchall mode class for modes which don’t fit into the above classes. Currently VOIDmode and BLKmode are in MODE_RANDOM.

Here are some C macros that relate to machine modes:

GET_MODE (x)
Returns the machine mode of the RTX x.

PUT_MODE (x, newmode)
Alters the machine mode of the RTX x to be newmode.

NUM_MACHINE_MODES
Stands for the number of machine modes available on the target machine. This is one greater than the largest numeric value of any machine mode.

GET_MODE_NAME (m)
Returns the name of mode m as a string.

GET_MODE_CLASS (m)
Returns the mode class of mode m.

GET_MODE_WIDER_MODE (m)
Returns the next wider natural mode. For example, the expression GET_MODE_WIDER_MODE (QImode) returns HImode.

GET_MODE_SIZE (m)
Returns the size in bytes of a datum of mode m.
GET_MODE_BITSIZE (m)
    Returns the size in bits of a datum of mode m.

GET_MODE_IBIT (m)
    Returns the number of integral bits of a datum of fixed-point mode m.

GET_MODE_FBIT (m)
    Returns the number of fractional bits of a datum of fixed-point mode m.

GET_MODE_MASK (m)
    Returns a bitmask containing 1 for all bits in a word that fit within mode m.
    This macro can only be used for modes whose bitsize is less than or equal to HOST_BITS_PER_INT.

GET_MODE_ALIGNMENT (m)
    Returns the required alignment, in bits, for an object of mode m.

GET_MODE_UNIT_SIZE (m)
    Returns the size in bytes of the subunits of a datum of mode m. This is the same as GET_MODE_SIZE except in the case of complex modes. For them, the unit size is the size of the real or imaginary part.

GET_MODE_NUNITS (m)
    Returns the number of units contained in a mode, i.e., GET_MODE_SIZE divided by GET_MODE_UNIT_SIZE.

GET_CLASS_NARROWEST_MODE (c)
    Returns the narrowest mode in mode class c.

    The global variables byte_mode and word_mode contain modes whose classes are MODE_INT and whose bitsizes are either BITS_PER_UNIT or BITS_PER_WORD, respectively. On 32-bit machines, these are QImode and SImode, respectively.

10.7 Constant Expression Types

The simplest RTL expressions are those that represent constant values.

(const_int i)
    This type of expression represents the integer value i. i is customarily accessed with the macro INTVAL as in INTVAL (exp), which is equivalent to XWINT (exp, 0).

    Constants generated for modes with fewer bits than HOST_WIDE_INT must be sign extended to full width (e.g., with gen_int_mode).

    There is only one expression object for the integer value zero; it is the value of the variable const0_rtx. Likewise, the only expression for integer value one is found in const1_rtx, the only expression for integer value two is found in const2_rtx, and the only expression for integer value negative one is found in constm1_rtx. Any attempt to create an expression of code const_int and value zero, one, two or negative one will return const0_rtx, const1_rtx, const2_rtx or constm1_rtx as appropriate.
Similarly, there is only one object for the integer whose value is \texttt{STORE\_FLAG\_VALUE}. It is found in \texttt{const\_true\_rtx}. If \texttt{STORE\_FLAG\_VALUE} is one, \texttt{const\_true\_rtx} and \texttt{const1\_rtx} will point to the same object. If \texttt{STORE\_FLAG\_VALUE} is \(-1\), \texttt{const\_true\_rtx} and \texttt{constm1\_rtx} will point to the same object.

\texttt{(const\_double:m i0 i1 ...)}

Represents either a floating-point constant of mode \texttt{m} or an integer constant too large to fit into \texttt{HOST\_BITS\_PER\_WIDE\_INT} bits but small enough to fit within twice that number of bits (GCC does not provide a mechanism to represent even larger constants). In the latter case, \texttt{m} will be \texttt{VOIDmode}.

If \texttt{m} is \texttt{VOIDmode}, the bits of the value are stored in \texttt{i0} and \texttt{i1}. \texttt{i0} is customarily accessed with the macro \texttt{CONST\_DOUBLE\_LOW} and \texttt{i1} with \texttt{CONST\_DOUBLE\_HIGH}.

If the constant is floating point (regardless of its precision), then the number of integers used to store the value depends on the size of \texttt{REAL\_VALUE\_TYPE} (see Section 17.23 [Floating Point], page 531). The integers represent a floating point number, but not precisely in the target machine’s or host machine’s floating point format. To convert them to the precise bit pattern used by the target machine, use the macro \texttt{REAL\_VALUE\_TO\_TARGET\_DOUBLE} and friends (see Section 17.21.2 [Data Output], page 500).

\texttt{(const\_fixed:m ...)}

Represents a fixed-point constant of mode \texttt{m}. The operand is a data structure of type \texttt{struct\ fixed\_value} and is accessed with the macro \texttt{CONST\_FIXED\_VALUE}. The high part of data is accessed with \texttt{CONST\_FIXED\_VALUE\_HIGH}; the low part is accessed with \texttt{CONST\_FIXED\_VALUE\_LOW}.

\texttt{(const\_vector:m [x0 x1 ...])}

Represents a vector constant. The square brackets stand for the vector containing the constant elements. \texttt{x0}, \texttt{x1} and so on are the \texttt{const\_int}, \texttt{const\_double} or \texttt{const\_fixed} elements.

The number of units in a \texttt{const\_vector} is obtained with the macro \texttt{CONST\_VECTOR\_NUNITS} as in \texttt{CONST\_VECTOR\_NUNITS (v)}.

Individual elements in a vector constant are accessed with the macro \texttt{CONST\_VECTOR\_ELT} as in \texttt{CONST\_VECTOR\_ELT (v, n)} where \texttt{v} is the vector constant and \texttt{n} is the element desired.

\texttt{(const\_string str)}

Represents a constant string with value \texttt{str}. Currently this is used only for insn attributes (see Section 16.19 [Insn Attributes], page 371) since constant strings in C are placed in memory.

\texttt{(symbol\_ref:mode symbol)}

Represents the value of an assembler label for data. \texttt{symbol} is a string that describes the name of the assembler label. If it starts with a ‘\*’, the label is the rest of \texttt{symbol} not including the ‘\*’. Otherwise, the label is \texttt{symbol}, usually prefixed with ‘\_’.

The \texttt{symbol\_ref} contains a mode, which is usually \texttt{Pmode}. Usually that is the only mode for which a symbol is directly valid.
(label_ref: mode label)
Represents the value of an assembler label for code. It contains one operand, an expression, which must be a code_label or a note of type NOTE_INSNN_DELETED_LABEL that appears in the instruction sequence to identify the place where the label should go.
The reason for using a distinct expression type for code label references is so that jump optimization can distinguish them.
The label_ref contains a mode, which is usually Pmode. Usually that is the only mode for which a label is directly valid.

(const:m exp)
Represents a constant that is the result of an assembly-time arithmetic computation. The operand, exp, is an expression that contains only constants (const_int, symbol_ref and label_ref expressions) combined with plus and minus. However, not all combinations are valid, since the assembler cannot do arbitrary arithmetic on relocatable symbols.
m should be Pmode.

(high:m exp)
Represents the high-order bits of exp, usually a symbol_ref. The number of bits is machine-dependent and is normally the number of bits specified in an instruction that initializes the high-order bits of a register. It is used with lo_sum to represent the typical two-instruction sequence used in RISC machines to reference a global memory location.
m should be Pmode.

The macro CONST0_RTX (mode) refers to an expression with value 0 in mode mode. If mode mode is of mode class MODE_INT, it returns const0_rtx. If mode mode is of mode class MODE_FLOAT, it returns a CONST_DOUBLE expression in mode mode. Otherwise, it returns a CONST_VECTOR expression in mode mode. Similarly, the macro CONST1_RTX (mode) refers to an expression with value 1 in mode mode and similarly for CONST2_RTX. The CONST1_RTX and CONST2_RTX macros are undefined for vector modes.

10.8 Registers and Memory
Here are the RTL expression types for describing access to machine registers and to main memory.

(reg:m n)
For small values of the integer n (those that are less than FIRST_PSEUDO_REGISTER), this stands for a reference to machine register number n: a hard register. For larger values of n, it stands for a temporary value or pseudo register. The compiler's strategy is to generate code assuming an unlimited number of such pseudo registers, and later convert them into hard registers or into memory references.
m is the machine mode of the reference. It is necessary because machines can generally refer to each register in more than one mode. For example, a register may contain a full word but there may be instructions to refer to it as a half
word or as a single byte, as well as instructions to refer to it as a floating point
number of various precisions.

Even for a register that the machine can access in only one mode, the mode
must always be specified.

The symbol FIRST_PSEUDO_REGISTER is defined by the machine description,
since the number of hard registers on the machine is an invariant characteristic
of the machine. Note, however, that not all of the machine registers must be
general registers. All the machine registers that can be used for storage of data
are given hard register numbers, even those that can be used only in certain
instructions or can hold only certain types of data.

A hard register may be accessed in various modes throughout one function,
but each pseudo register is given a natural mode and is accessed only in that
mode. When it is necessary to describe an access to a pseudo register using a
nonnatural mode, a subreg expression is used.

A reg expression with a machine mode that specifies more than one word
of data may actually stand for several consecutive registers. If in addition the
register number specifies a hardware register, then it actually represents several
consecutive hardware registers starting with the specified one.

Each pseudo register number used in a function’s RTL code is represented by
a unique reg expression.

Some pseudo register numbers, those within the range of FIRST_VIRTUAL_
REGISTER to LAST_VIRTUAL_REGISTER only appear during the RTL generation
phase and are eliminated before the optimization phases. These represent loca-
tions in the stack frame that cannot be determined until RTL generation for
the function has been completed. The following virtual register numbers are
defined:

VIRTUAL_INCOMING_ARGS_REGNUM
This points to the first word of the incoming arguments passed
on the stack. Normally these arguments are placed there by the
caller, but the callee may have pushed some arguments that were
previously passed in registers.

When RTL generation is complete, this virtual register is replaced
by the sum of the register given by ARG_POINTER_REGNUM and the
value of FIRST_PARM_OFFSET.

VIRTUAL_STACK_VARS_REGNUM
If FRAME_GROWS_DOWNWARD is defined to a nonzero value, this points
to immediately above the first variable on the stack. Otherwise, it
points to the first variable on the stack.

VIRTUAL_STACK_VARS_REGNUM is replaced with the sum of the reg-
ister given by FRAME_POINTER_REGNUM and the value STARTING_ FRAME_OFFSET.
VIRTUAL_STACK_DYNAMIC_REGNUM
This points to the location of dynamically allocated memory on the
stack immediately after the stack pointer has been adjusted by the
amount of memory desired.
This virtual register is replaced by the sum of the register given by
STACK_POINTER_REGNUM and the value STACK_DYNAMIC_OFFSET.

VIRTUAL_OUTGOING_ARGS_REGNUM
This points to the location in the stack at which outgoing arguments
should be written when the stack is pre-pushed (arguments pushed
using push insns should always use STACK_POINTER_REGNUM).
This virtual register is replaced by the sum of the register given by
STACK_POINTER_REGNUM and the value STACK_POINTER_OFFSET.

(subreg:m1 reg:m2 bytenum)

subreg expressions are used to refer to a register in a machine mode other than
its natural one, or to refer to one register of a multi-part reg that actually refers
to several registers.
Each pseudo register has a natural mode. If it is necessary to operate on it in
a different mode, the register must be enclosed in a subreg.

There are currently three supported types for the first operand of a subreg:

• pseudo registers This is the most common case. Most subregs have pseudo
  regs as their first operand.
• mem subregs of mem were common in earlier versions of GCC and are still
  supported. During the reload pass these are replaced by plain mems. On
  machines that do not do instruction scheduling, use of subregs of mem are
  still used, but this is no longer recommended. Such subregs are considered
to be register_operands rather than memory_operands before and during
  reload. Because of this, the scheduling passes cannot properly schedule
  instructions with subreg operands of mem, so for machines that do scheduling,
  subregs of mem should never be used. To support this, the combine and
  recog passes have explicit code to inhibit the creation of subregs of mem
  when INSN_SCHEDULING is defined.
  The use of subregs of mem after the reload pass is an area that is not well
  understood and should be avoided. There is still some code in the compiler
to support this, but this code has possibly rotted. This use of subregs is
discouraged and will most likely not be supported in the future.
• hard registers It is seldom necessary to wrap hard registers in subregs; such
  registers would normally reduce to a single reg rtx. This use of subregs is
discouraged and may not be supported in the future.

subregs of subregs are not supported. Using simplify_gen_subreg is the
recommended way to avoid this problem.

subregs come in two distinct flavors, each having its own usage and rules:

Paradoxical subregs
When m1 is strictly wider than m2, the subreg expression is called
paradoxical. The canonical test for this class of subreg is:
GET_MODE_SIZE (m1) > GET_MODE_SIZE (m2)

Paradoxical subregs can be used as both lvalues and rvalues. When used as an lvalue, the low-order bits of the source value are stored in reg and the high-order bits are discarded. When used as an rvalue, the low-order bits of the subreg are taken from reg while the high-order bits may or may not be defined.

The high-order bits of rvalues are in the following circumstances:

- subregs of mem When m2 is smaller than a word, the macro LOAD_EXTEND_OP, can control how the high-order bits are defined.

- subreg of regs The upper bits are defined when SUBREG_PROMOTED_VAR_P is true. SUBREG_PROMOTED_UNSIGNED_P describes what the upper bits hold. Such subregs usually represent local variables, register variables and parameter pseudo variables that have been promoted to a wider mode.

bytenum is always zero for a paradoxical subreg, even on big-endian targets.

For example, the paradoxical subreg:

(set (subreg:SI (reg:HI x) 0) y)

stores the lower 2 bytes of y in x and discards the upper 2 bytes. A subsequent:

(set z (subreg:SI (reg:HI x) 0))

would set the lower two bytes of z to y and set the upper two bytes to an unknown value assuming SUBREG_PROMOTED_VAR_P is false.

Normal subregs

When m1 is at least as narrow as m2 the subreg expression is called normal.

Normal subregs restrict consideration to certain bits of reg. There are two cases. If m1 is smaller than a word, the subreg refers to the least-significant part (or lowpart) of one word of reg. If m1 is word-sized or greater, the subreg refers to one or more complete words.

When used as an lvalue, subreg is a word-based accessor. Storing to a subreg modifies all the words of reg that overlap the subreg, but it leaves the other words of reg alone.

When storing to a normal subreg that is smaller than a word, the other bits of the referenced word are usually left in an undefined state. This laxity makes it easier to generate efficient code for such instructions. To represent an instruction that preserves all the bits outside of those in the subreg, use strict_low_part or zero_extract around the subreg.

bytenum must identify the offset of the first byte of the subreg from the start of reg, assuming that reg is laid out in memory.
order. The memory order of bytes is defined by two target macros, 
WORDS_BIG_ENDIAN and BYTES_BIG_ENDIAN:

- **WORDS_BIG_ENDIAN**, if set to 1, says that byte number zero is 
  part of the most significant word; otherwise, it is part of the 
  least significant word.

- **BYTES_BIG_ENDIAN**, if set to 1, says that byte number zero is 
  the most significant byte within a word; otherwise, it is the 
  least significant byte within a word.

On a few targets, FLOAT_WORDS_BIG_ENDIAN disagrees with WORDS_ 
BIG_ENDIAN. However, most parts of the compiler treat floating 
point values as if they had the same endianness as integer values. 
This works because they handle them solely as a collection of integer 
values, with no particular numerical value. Only real.c and the 
runtime libraries care about FLOAT_WORDS_BIG_ENDIAN.

Thus, 

```
(subreg:HI (reg:SI x) 2)
```

on a BYTES_BIG_ENDIAN, ‘UNITS_PER_WORD == 4’ target is the same 
as 

```
(subreg:HI (reg:SI x) 0)
```

on a little-endian, ‘UNITS_PER_WORD == 4’ target. Both subregs 
access the lower two bytes of register x.

A **MODE_PARTIAL_INT** mode behaves as if it were as wide as the corresponding 
**MODE_INT** mode, except that it has an unknown number of undefined bits. For 
example:

```
(subreg:PSI (reg:SI 0) 0)
```

accesses the whole of ‘(reg:SI 0)’, but the exact relationship between 
the PSImode value and the SImode value is not defined. If we assume 
‘UNITS_PER_WORD <= 4’, then the following two subregs:

```
(subreg:PSI (reg:DI 0) 0)
(subreg:PSI (reg:DI 0) 4)
```

represent independent 4-byte accesses to the two halves of ‘(reg:DI 0)’. Both 
subregs have an unknown number of undefined bits.

If ‘UNITS_PER_WORD <= 2’ then these two subregs:

```
(subreg:HI (reg:PSI 0) 0)
(subreg:HI (reg:PSI 0) 2)
```

represent independent 2-byte accesses that together span the whole of 
‘(reg:PSI 0)’. Storing to the first subreg does not affect the value of the 
second, and vice versa. ‘(reg:PSI 0)’ has an unknown number of undefined 
bits, so the assignment:

```
(set (subreg:HI (reg:PSI 0) 0) (reg:HI 4))
```

does not guarantee that ‘(subreg:HI (reg:PSI 0) 0)’ has the value ‘(reg:HI 
4)’.

The rules above apply to both pseudo regs and hard regs. If the semantics 
are not correct for particular combinations of m1, m2 and hard reg, the target-
specific code must ensure that those combinations are never used. For example:
CANNOT_CHANGE_MODE_CLASS (m2, m1, class)
must be true for every class class that includes reg.
The first operand of a subreg expression is customarily accessed with the
SUBREG_REG macro and the second operand is customarily accessed with the
SUBREG_BYTE macro.
It has been several years since a platform in which BYTES_BIG_ENDIAN not
equal to WORDS_BIG_ENDIAN has been tested. Anyone wishing to support such
a platform in the future may be confronted with code rot.

(scratch:m)
This represents a scratch register that will be required for the execution of a
single instruction and not used subsequently. It is converted into a reg by either
the local register allocator or the reload pass.
scratch is usually present inside a clobber operation (see Section 10.15 [Side
Effects], page 159).

(cc0)
This refers to the machine’s condition code register. It has no operands and
may not have a machine mode. There are two ways to use it:
• To stand for a complete set of condition code flags. This is best on most
machines, where each comparison sets the entire series of flags.
  With this technique, (cc0) may be validly used in only two contexts: as
  the destination of an assignment (in test and compare instructions) and in
  comparison operators comparing against zero (const_int with value zero;
  that is to say, const0_rtx).
• To stand for a single flag that is the result of a single condition. This is
  useful on machines that have only a single flag bit, and in which comparison
  instructions must specify the condition to test.
  With this technique, (cc0) may be validly used in only two contexts: as the
  destination of an assignment (in test and compare instructions) where the
  source is a comparison operator, and as the first operand of if_then_else
  (in a conditional branch).
There is only one expression object of code cc0; it is the value of the variable
cc0_rtx. Any attempt to create an expression of code cc0 will return cc0_rtx.
Instructions can set the condition code implicitly. On many machines, nearly
all instructions set the condition code based on the value that they compute or
store. It is not necessary to record these actions explicitly in the RTL because
the machine description includes a prescription for recognizing the instructions
that do so (by means of the macro NOTICE_UPDATE_CC). See Section 17.16
[Condition Code], page 476. Only instructions whose sole purpose is to set
the condition code, and instructions that use the condition code, need mention
(cc0).
On some machines, the condition code register is given a register number and
a reg is used instead of (cc0). This is usually the preferable approach if only
a small subset of instructions modify the condition code. Other machines store
condition codes in general registers; in such cases a pseudo register should be
used.
Some machines, such as the SPARC and RS/6000, have two sets of arithmetic instructions, one that sets and one that does not set the condition code. This is best handled by normally generating the instruction that does not set the condition code, and making a pattern that both performs the arithmetic and sets the condition code register (which would not be (cc0) in this case). For examples, search for ‘addcc’ and ‘andcc’ in ‘sparc.md’.

This represents the machine’s program counter. It has no operands and may not have a machine mode. (pc) may be validly used only in certain specific contexts in jump instructions.

There is only one expression object of code pc; it is the value of the variable pc_rtx. Any attempt to create an expression of code pc will return pc_rtx.

All instructions that do not jump alter the program counter implicitly by incrementing it, but there is no need to mention this in the RTL.

This RTX represents a reference to main memory at an address represented by the expression addr. m specifies how large a unit of memory is accessed. alias specifies an alias set for the reference. In general two items are in different alias sets if they cannot reference the same memory address.

The construct (mem:BLK (scratch)) is considered to alias all other memories. Thus it may be used as a memory barrier in epilogue stack deallocation patterns.

This RTX represents the concatenation of two other RTXs. This is used for complex values. It should only appear in the RTL attached to declarations and during RTL generation. It should not appear in the ordinary insn chain.

This RTX represents the concatenation of all the rtx to make a single value. Like concat, this should only appear in declarations, and not in the insn chain.

10.9 RTL Expressions for Arithmetic

Unless otherwise specified, all the operands of arithmetic expressions must be valid for mode m. An operand is valid for mode m if it has mode m, or if it is a const_int or const_double and m is a mode of class MODE_INT.

For commutative binary operations, constants should be placed in the second operand.

These three expressions all represent the sum of the values represented by x and y carried out in machine mode m. They differ in their behavior on overflow of integer modes. plus wraps round modulo the width of m; ss_plus saturates at the maximum signed value representable in m; us_plus saturates at the maximum unsigned value.

This expression represents the sum of x and the low-order bits of y. It is used with high (see Section 10.7 [Constants], page 142) to represent the typical
two-instruction sequence used in RISC machines to reference a global memory location.
The number of low order bits is machine-dependent but is normally the number of bits in a Pmode item minus the number of bits set by high. $m$ should be Pmode.

\[
\text{(minus:} \ mode \ x \ y) \\
\text{(ss_minus:} \ mode \ x \ y) \\
\text{(us_minus:} \ mode \ x \ y)
\]
These three expressions represent the result of subtracting $y$ from $x$, carried out in mode $M$. Behavior on overflow is the same as for the three variants of plus (see above).

\[
\text{(compare:} \ mode \ x \ y)
\]
Represents the result of subtracting $y$ from $x$ for purposes of comparison. The result is computed without overflow, as if with infinite precision.

Of course, machines can't really subtract with infinite precision. However, they can pretend to do so when only the sign of the result will be used, which is the case when the result is stored in the condition code. And that is the only way this kind of expression may validly be used: as a value to be stored in the condition codes, either (cc0) or a register. See Section 10.10 [Comparisons], page 154.

The mode $m$ is not related to the modes of $x$ and $y$, but instead is the mode of the condition code value. If (cc0) is used, it is VOIDmode. Otherwise it is some mode in class MODE_CC, often CCmode. See Section 17.16 [Condition Code], page 476. If $m$ is VOIDmode or CCmode, the operation returns sufficient information (in an unspecified format) so that any comparison operator can be applied to the result of the COMPARE operation. For other modes in class MODE_CC, the operation only returns a subset of this information.

Normally, $x$ and $y$ must have the same mode. Otherwise, compare is valid only if the mode of $x$ is in class MODE_INT and $y$ is a \texttt{const_int} or \texttt{const_double} with mode VOIDmode. The mode of $x$ determines what mode the comparison is to be done in; thus it must not be VOIDmode.

If one of the operands is a constant, it should be placed in the second operand and the comparison code adjusted as appropriate.

A compare specifying two VOIDmode constants is not valid since there is no way to know in what mode the comparison is to be performed; the comparison must either be folded during the compilation or the first operand must be loaded into a register while its mode is still known.

\[
\text{(neg:} \ mode \ x) \\
\text{(ss_neg:} \ mode \ x) \\
\text{(us_neg:} \ mode \ x)
\]
These two expressions represent the negation (subtraction from zero) of the value represented by $x$, carried out in mode $m$. They differ in the behavior on overflow of integer modes. In the case of neg, the negation of the operand may be a number not representable in mode $m$, in which case it is truncated.
to \( m \). \texttt{ss\_neg} and \texttt{us\_neg} ensure that an out-of-bounds result saturates to the maximum or minimum signed or unsigned value.

\( \text{(mult:\!/m x y)} \)
\( \text{(ss\_mult:\!/m x y)} \)
\( \text{(us\_mult:\!/m x y)} \)

Represents the signed product of the values represented by \( x \) and \( y \) carried out in machine mode \( m \). \texttt{ss\_mult} and \texttt{us\_mult} ensure that an out-of-bounds result saturates to the maximum or minimum signed or unsigned value.

Some machines support a multiplication that generates a product wider than the operands. Write the pattern for this as

\( \text{(mult:\!/m (sign\_extend:\!/m x) (sign\_extend:\!/m y))} \)

where \( m \) is wider than the modes of \( x \) and \( y \), which need not be the same.

For unsigned widening multiplication, use the same idiom, but with \texttt{zero\_extend} instead of \texttt{sign\_extend}.

\( \text{(fma:\!/m x y z)} \)

Represents the \texttt{fma}, \texttt{fmaf}, and \texttt{fmal} builtin functions that do a combined multiply of \( x \) and \( y \) and then adding to \( z \) without doing an intermediate rounding step.

\( \text{(div:\!/m x y)} \)
\( \text{(ss\_div:\!/m x y)} \)

Represents the quotient in signed division of \( x \) by \( y \), carried out in machine mode \( m \). If \( m \) is a floating point mode, it represents the exact quotient; otherwise, the integerized quotient. \texttt{ss\_div} ensures that an out-of-bounds result saturates to the maximum or minimum signed value.

Some machines have division instructions in which the operands and quotient widths are not all the same; you should represent such instructions using \texttt{truncate} and \texttt{sign\_extend} as in,

\( \text{(truncate:\!/m1 (div:\!/m2 x (sign\_extend:\!/m2 y)))} \)

\( \text{(udiv:\!/m x y)} \)
\( \text{(us\_div:\!/m x y)} \)

Like \texttt{div} but represents unsigned division. \texttt{us\_div} ensures that an out-of-bounds result saturates to the maximum or minimum unsigned value.

\( \text{(mod:\!/m x y)} \)
\( \text{(umod:\!/m x y)} \)

Like \texttt{div} and \texttt{udiv} but represent the remainder instead of the quotient.

\( \text{(smin:\!/m x y)} \)
\( \text{(smax:\!/m x y)} \)

Represents the smaller (for \texttt{smin}) or larger (for \texttt{smax}) of \( x \) and \( y \), interpreted as signed values in mode \( m \). When used with floating point, if both operands are zeros, or if either operand is NaN, then it is unspecified which of the two operands is returned as the result.

\( \text{(umin:\!/m x y)} \)
\( \text{(umax:\!/m x y)} \)

Like \texttt{smin} and \texttt{smax}, but the values are interpreted as unsigned integers.
(not:m x)
    Represents the bitwise complement of the value represented by x, carried out in mode m, which must be a fixed-point machine mode.

(and:m x y)
    Represents the bitwise logical-and of the values represented by x and y, carried out in machine mode m, which must be a fixed-point machine mode.

(ior:m x y)
    Represents the bitwise inclusive-or of the values represented by x and y, carried out in machine mode m, which must be a fixed-point mode.

(xor:m x y)
    Represents the bitwise exclusive-or of the values represented by x and y, carried out in machine mode m, which must be a fixed-point mode.

(ashift:m x c)
(ssh_ashift:m x c)
(us_ashift:m x c)
    These three expressions represent the result of arithmetically shifting x left by c places. They differ in their behavior on overflow of integer modes. An ashift operation is a plain shift with no special behavior in case of a change in the sign bit; ss_ashift and us_ashift saturates to the minimum or maximum representable value if any of the bits shifted out differs from the final sign bit. x have mode m, a fixed-point machine mode. c be a fixed-point mode or be a constant with mode VOIDmode; which mode is determined by the mode called for in the machine description entry for the left-shift instruction. For example, on the VAX, the mode of c is QImode regardless of m.

(lshiftrt:m x c)
(ashiftrt:m x c)
    Like ashift but for right shift. Unlike the case for left shift, these two operations are distinct.

(rotate:m x c)
(rotatert:m x c)
    Similar but represent left and right rotate. If c is a constant, use rotate.

(abs:m x)
(ss_abs:m x)
    Represents the absolute value of x, computed in mode m. ss_abs ensures that an out-of-bounds result saturates to the maximum signed value.

(sqrt:m x)
    Represents the square root of x, computed in mode m. Most often m will be a floating point mode.

(ffs:m x)
    Represents one plus the index of the least significant 1-bit in x, represented as an integer of mode m. (The value is zero if x is zero.) The mode of x need not be m; depending on the target machine, various mode combinations may be valid.
Represents the number of leading 0-bits in \( x \), represented as an integer of mode \( m \), starting at the most significant bit position. If \( x \) is zero, the value is determined by \texttt{CLZ_DEFINED_VALUE_AT_ZERO} (see Section 17.31 [Misc], page 542). Note that this is one of the few expressions that is not invariant under widening. The mode of \( x \) will usually be an integer mode.

(bswap: \( m \) \( x \))

Represents the value \( x \) with the order of bytes reversed, carried out in mode \( m \), which must be a fixed-point machine mode.

### 10.10 Comparison Operations

Comparison operators test a relation on two operands and are considered to represent a machine-dependent nonzero value described by, but not necessarily equal to, \texttt{STORE_FLAG_VALUE} (see Section 17.31 [Misc], page 542) if the relation holds, or zero if it does not, for comparison operators whose results have a ‘MODE_INT’ mode, \texttt{FLOAT_STORE_FLAG_VALUE} (see Section 17.31 [Misc], page 542) if the relation holds, or zero if it does not, for comparison operators that return floating-point values, and a vector of either \texttt{VECTOR_STORE_FLAG_VALUE} (see Section 17.31 [Misc], page 542) if the relation holds, or of zeros if it does not, for comparison operators that return vector results. The mode of the comparison operation is independent of the mode of the data being compared. If the comparison operation is being tested (e.g., the first operand of an \texttt{if.then.else}), the mode must be \texttt{VOIDmode}.

There are two ways that comparison operations may be used. The comparison operators may be used to compare the condition codes (cc0) against zero, as in \texttt{(eq (cc0) (const_int 0))}. Such a construct actually refers to the result of the preceding instruction in which the condition codes were set. The instruction setting the condition code must be adjacent to the instruction using the condition code; only \texttt{note} insns may separate them.

Alternatively, a comparison operation may directly compare two data objects. The mode of the comparison is determined by the operands; they must both be valid for a common machine mode. A comparison with both operands constant would be invalid as the machine mode could not be deduced from it, but such a comparison should never exist in RTL due to constant folding.
In the example above, if \((\text{cc}0)\) were last set to \((\text{compare } x \ y)\), the comparison operation is identical to \((\text{eq } x \ y)\). Usually only one style of comparisons is supported on a particular machine, but the combine pass will try to merge the operations to produce the \text{eq} shown in case it exists in the context of the particular insn involved.

Inequality comparisons come in two flavors, signed and unsigned. Thus, there are distinct expression codes \text{gt} and \text{gtu} for signed and unsigned greater-than. These can produce different results for the same pair of integer values: for example, 1 is signed greater-than \(-1\) but not unsigned greater-than, because \(-1\) when regarded as unsigned is actually \(0xffffffff\) which is greater than 1.

The signed comparisons are also used for floating point values. Floating point comparisons are distinguished by the machine modes of the operands.

\[
\begin{align*}
\text{(eq:} & m \ x \ y) \\
& \text{STORE_FLAG_VALUE if the values represented by } x \text{ and } y \text{ are equal, otherwise } 0. \\
\text{(ne:} & m \ x \ y) \\
& \text{STORE_FLAG_VALUE if the values represented by } x \text{ and } y \text{ are not equal, otherwise } 0. \\
\text{(gt:} & m \ x \ y) \\
& \text{STORE_FLAG_VALUE if the } x \text{ is greater than } y. \text{ If they are fixed-point, the comparison is done in a signed sense.} \\
\text{(gtu:} & m \ x \ y) \\
& \text{Like } \text{gt} \text{ but does unsigned comparison, on fixed-point numbers only.} \\
\text{(lt:} & m \ x \ y) \\
\text{(ltu:} & m \ x \ y) \\
& \text{Like } \text{gt} \text{ and } \text{gtu} \text{ but test for “less than”.} \\
\text{(ge:} & m \ x \ y) \\
\text{(geu:} & m \ x \ y) \\
& \text{Like } \text{gt} \text{ and } \text{gtu} \text{ but test for “greater than or equal”.} \\
\text{(le:} & m \ x \ y) \\
\text{(leu:} & m \ x \ y) \\
& \text{Like } \text{gt} \text{ and } \text{gtu} \text{ but test for “less than or equal”.} \\
\text{(if\_then\_else cond then else)} \\
& \text{This is not a comparison operation but is listed here because it is always used in conjunction with a comparison operation. To be precise, } \text{cond} \text{ is a comparison expression. This expression represents a choice, according to } \text{cond}, \text{ between the value represented by } \text{then} \text{ and the one represented by } \text{else}. \\
& \text{On most machines, } \text{if\_then\_else} \text{ expressions are valid only to express conditional jumps.} \\
\text{(cond [test1 value1 test2 value2 ...] default)} \\
& \text{Similar to } \text{if\_then\_else}, \text{ but more general. Each of } \text{test1}, \text{ test2}, \text{ ... is performed in turn. The result of this expression is the value corresponding to the first nonzero test, or } \text{default} \text{ if none of the tests are nonzero expressions.} \\
& \text{This is currently not valid for instruction patterns and is supported only for insn attributes. See Section 16.19 [Insn Attributes], page 371.} 
\end{align*}
\]
10.11 Bit-Fields

Special expression codes exist to represent bit-field instructions.

(sign_extract:m loc size pos)

This represents a reference to a sign-extended bit-field contained or starting in loc (a memory or register reference). The bit-field is size bits wide and starts at bit pos. The compilation option BITS_BIG_ENDIAN says which end of the memory unit pos counts from.

If loc is in memory, its mode must be a single-byte integer mode. If loc is in a register, the mode to use is specified by the operand of the insv or extv pattern (see Section 16.9 [Standard Names], page 332) and is usually a full-word integer mode, which is the default if none is specified.

The mode of pos is machine-specific and is also specified in the insv or extv pattern.

The mode m is the same as the mode that would be used for loc if it were a register.

A sign_extract can not appear as an lvalue, or part thereof, in RTL.

(zero_extract:m loc size pos)

Like sign_extract but refers to an unsigned or zero-extended bit-field. The same sequence of bits are extracted, but they are filled to an entire word with zeros instead of by sign-extension.

Unlike sign_extract, this type of expressions can be lvalues in RTL; they may appear on the left side of an assignment, indicating insertion of a value into the specified bit-field.

10.12 Vector Operations

All normal RTL expressions can be used with vector modes; they are interpreted as operating on each part of the vector independently. Additionally, there are a few new expressions to describe specific vector operations.

(vec_merge:m vec1 vec2 items)

This describes a merge operation between two vectors. The result is a vector of mode m; its elements are selected from either vec1 or vec2. Which elements are selected is described by items, which is a bit mask represented by a const_int; a zero bit indicates the corresponding element in the result vector is taken from vec2 while a set bit indicates it is taken from vec1.

(vec_select:m vec1 selection)

This describes an operation that selects parts of a vector. vec1 is the source vector, and selection is a parallel that contains a const_int for each of the subparts of the result vector, giving the number of the source subpart that should be stored into it. The result mode m is either the submode for a single element of vec1 (if only one subpart is selected), or another vector mode with that element submode (if multiple subparts are selected).
(vec_concat: m vec1 vec2)
  Describes a vector concat operation. The result is a concatenation of the vectors vec1 and vec2; its length is the sum of the lengths of the two inputs.

(vec_duplicate: m vec)
  This operation converts a small vector into a larger one by duplicating the input values. The output vector mode must have the same submodes as the input vector mode, and the number of output parts must be an integer multiple of the number of input parts.

10.13 Conversions

All conversions between machine modes must be represented by explicit conversion operations. For example, an expression which is the sum of a byte and a full word cannot be written as (plus: SI (reg: QI 34) (reg: SI 80)) because the plus operation requires two operands of the same machine mode. Therefore, the byte-sized operand is enclosed in a conversion operation, as in

  (plus: SI (sign_extend: SI (reg: QI 34)) (reg: SI 80))

The conversion operation is not a mere placeholder, because there may be more than one way of converting from a given starting mode to the desired final mode. The conversion operation code says how to do it.

For all conversion operations, x must not be VOIDmode because the mode in which to do the conversion would not be known. The conversion must either be done at compile-time or x must be placed into a register.

(sign_extend: m x)
  Represents the result of sign-extending the value x to machine mode m. m must be a fixed-point mode and x a fixed-point value of a mode narrower than m.

(zero_extend: m x)
  Represents the result of zero-extending the value x to machine mode m. m must be a fixed-point mode and x a fixed-point value of a mode narrower than m.

(float_extend: m x)
  Represents the result of extending the value x to machine mode m. m must be a floating point mode and x a floating point value of a mode narrower than m.

(truncate: m x)
  Represents the result of truncating the value x to machine mode m. m must be a fixed-point mode and x a fixed-point value of a mode wider than m.

(ss_truncate: m x)
  Represents the result of truncating the value x to machine mode m, using signed saturation in the case of overflow. Both m and the mode of x must be fixed-point modes.

(us_truncate: m x)
  Represents the result of truncating the value x to machine mode m, using unsigned saturation in the case of overflow. Both m and the mode of x must be fixed-point modes.
(float_truncate::m x)
Represents the result of truncating the value x to machine mode m. m must be a floating point mode and x a floating point value of a mode wider than m.

(float::m x)
Represents the result of converting fixed point value x, regarded as signed, to floating point mode m.

(unsigned_float::m x)
Represents the result of converting fixed point value x, regarded as unsigned, to floating point mode m.

(fix::m x)
When m is a floating-point mode, represents the result of converting floating point value x (valid for mode m) to an integer, still represented in floating point mode m, by rounding towards zero.
When m is a fixed-point mode, represents the result of converting floating point value x to mode m, regarded as signed. How rounding is done is not specified, so this operation may be used validly in compiling C code only for integer-valued operands.

(unsigned_fix::m x)
Represents the result of converting floating point value x to fixed point mode m, regarded as unsigned. How rounding is done is not specified.

(frac_convert::m x)
Represents the result of converting fixed-point value x to fixed-point mode m, signed integer value x to fixed-point mode m, floating-point value x to fixed-point mode m, fixed-point value x to integer mode m regarded as signed, or fixed-point value x to floating-point mode m. When overflows or underflows happen, the results are undefined.

(sat_frac::m x)
Represents the result of converting fixed-point value x to fixed-point mode m, signed integer value x to fixed-point mode m, or floating-point value x to fixed-point mode m. When overflows or underflows happen, the results are saturated to the maximum or the minimum.

(unsigned_frac_convert::m x)
Represents the result of converting fixed-point value x to integer mode m regarded as unsigned, or unsigned integer value x to fixed-point mode m. When overflows or underflows happen, the results are undefined.

(unsigned_sat_frac::m x)
Represents the result of converting unsigned integer value x to fixed-point mode m. When overflows or underflows happen, the results are saturated to the maximum or the minimum.

10.14 Declarations
Declaration expression codes do not represent arithmetic operations but rather state assertions about their operands.
This expression code is used in only one context: as the destination operand of a set expression. In addition, the operand of this expression must be a non-paradoxical subreg expression.

The presence of strict_low_part says that the part of the register which is meaningful in mode \( n \), but is not part of mode \( m \), is not to be altered. Normally, an assignment to such a subreg is allowed to have undefined effects on the rest of the register when \( m \) is less than a word.

10.15 Side Effect Expressions

The expression codes described so far represent values, not actions. But machine instructions never produce values; they are meaningful only for their side effects on the state of the machine. Special expression codes are used to represent side effects.

The body of an instruction is always one of these side effect codes; the codes described above, which represent values, appear only as the operands of these.

\((\text{set } lval \ x)\)

Represents the action of storing the value of \( x \) into the place represented by \( lval \). \( lval \) must be an expression representing a place that can be stored in: reg (or subreg, strict_low_part or zero_extract), mem, pc, parallel, or cc0.

If \( lval \) is a reg, subreg or mem, it has a machine mode; then \( x \) must be valid for that mode.

If \( lval \) is a reg whose machine mode is less than the full width of the register, then it means that the part of the register specified by the machine mode is given the specified value and the rest of the register receives an undefined value. Likewise, if \( lval \) is a subreg whose machine mode is narrower than the mode of the register, the rest of the register can be changed in an undefined way.

If \( lval \) is a strict_low_part of a subreg, then the part of the register specified by the machine mode of the subreg is given the value \( x \) and the rest of the register is not changed.

If \( lval \) is a zero_extract, then the referenced part of the bit-field (a memory or register reference) specified by the zero_extract is given the value \( x \) and the rest of the bit-field is not changed. Note that sign_extract can not appear in \( lval \).

If \( lval \) is (cc0), it has no machine mode, and \( x \) may be either a compare expression or a value that may have any mode. The latter case represents a “test” instruction. The expression \((\text{set } (cc0) (\text{reg:m } n))\) is equivalent to \((\text{set } (cc0) (\text{compare } (\text{reg:m } n) (\text{const_int 0})))\). Use the former expression to save space during the compilation.

If \( lval \) is a parallel, it is used to represent the case of a function returning a structure in multiple registers. Each element of the parallel is an expr_list whose first operand is a reg and whose second operand is a const_int representing the offset (in bytes) into the structure at which the data in that register corresponds. The first element may be null to indicate that the structure is also passed partly in memory.
If `lval` is `(pc)`, we have a jump instruction, and the possibilities for `x` are very limited. It may be a `label_ref` expression (unconditional jump). It may be an `if_then_else` (conditional jump), in which case either the second or the third operand must be `(pc)` (for the case which does not jump) and the other of the two must be a `label_ref` (for the case which does jump). `x` may also be a `mem` or `(plus:SI (pc) y)`, where `y` may be a `reg` or a `mem`; these unusual patterns are used to represent jumps through branch tables.

If `lval` is neither `(cc0)` nor `(pc)`, the mode of `lval` must not be `VOIDmode` and the mode of `x` must be valid for the mode of `lval`.

`lval` is customarily accessed with the `SET_DEST` macro and `x` with the `SET_SRC` macro.

**(return)** As the sole expression in a pattern, represents a return from the current function, on machines where this can be done with one instruction, such as VAXen. On machines where a multi-instruction “epilogue” must be executed in order to return from the function, returning is done by jumping to a label which precedes the epilogue, and the `return` expression code is never used.

Inside an `if_then_else` expression, represents the value to be placed in `pc` to return to the caller.

Note that an insn pattern of (return) is logically equivalent to (set (pc) (return)), but the latter form is never used.

**(call function nargs)**

Represents a function call. `function` is a `mem` expression whose address is the address of the function to be called. `nargs` is an expression which can be used for two purposes: on some machines it represents the number of bytes of stack argument; on others, it represents the number of argument registers.

Each machine has a standard machine mode which `function` must have. The machine description defines macro `FUNCTION_MODE` to expand into the requisite mode name. The purpose of this mode is to specify what kind of addressing is allowed, on machines where the allowed kinds of addressing depend on the machine mode being addressed.

**(clobber x)**

Represents the storing or possible storing of an unpredictable, undescribed value into `x`, which must be a `reg`, `scratch`, `parallel` or `mem` expression.

One place this is used is in string instructions that store standard values into particular hard registers. It may not be worth the trouble to describe the values that are stored, but it is essential to inform the compiler that the registers will be altered, lest it attempt to keep data in them across the string instruction.

If `x` is `(mem:BLK (const_int 0))` or `(mem:BLK (scratch))`, it means that all memory locations must be presumed clobbered. If `x` is a `parallel`, it has the same meaning as a `parallel` in a `set` expression.

Note that the machine description classifies certain hard registers as “call-clobbered”. All function call instructions are assumed by default to clobber these registers, so there is no need to use `clobber` expressions to indicate this
fact. Also, each function call is assumed to have the potential to alter any memory location, unless the function is declared \texttt{const}.

If the last group of expressions in a \texttt{parallel} are each a \texttt{clobber} expression whose arguments are \texttt{reg} or \texttt{match\_scratch} (see Section 16.4 [RTL Template], page 291) expressions, the combiner phase can add the appropriate \texttt{clobber} expressions to an insn it has constructed when doing so will cause a pattern to be matched.

This feature can be used, for example, on a machine that whose multiply and add instructions don’t use an MQ register but which has an add-accumulate instruction that does clobber the MQ register. Similarly, a combined instruction might require a temporary register while the constituent instructions might not.

When a \texttt{clobber} expression for a register appears inside a \texttt{parallel} with other side effects, the register allocator guarantees that the register is unoccupied both before and after that insn if it is a hard register clobber. For pseudo-register clobber, the register allocator and the reload pass do not assign the same hard register to the clobber and the input operands if there is an insn alternative containing the ‘&’ constraint (see Section 16.8.4 [Modifiers], page 307) for the clobber and the hard register is in register classes of the clobber in the alternative. You can clobber either a specific hard register, a pseudo register, or a \texttt{scratch} expression; in the latter two cases, GCC will allocate a hard register that is available there for use as a temporary.

For instructions that require a temporary register, you should use \texttt{scratch} instead of a pseudo-register because this will allow the combiner phase to add the \texttt{clobber} when required. You do this by coding (\texttt{clobber (match\_scratch \ldots)}). If you do clobber a pseudo register, use one which appears nowhere else—generate a new one each time. Otherwise, you may confuse CSE.

There is one other known use for clobbering a pseudo register in a \texttt{parallel}: when one of the input operands of the insn is also clobbered by the insn. In this case, using the same pseudo register in the clobber and elsewhere in the insn produces the expected results.

\texttt{(use x)} Represents the use of the value of \texttt{x}. It indicates that the value in \texttt{x} at this point in the program is needed, even though it may not be apparent why this is so. Therefore, the compiler will not attempt to delete previous instructions whose only effect is to store a value in \texttt{x}. \texttt{x} must be a \texttt{reg} expression.

In some situations, it may be tempting to add a \texttt{use} of a register in a \texttt{parallel} to describe a situation where the value of a special register will modify the behavior of the instruction. A hypothetical example might be a pattern for an addition that can either wrap around or use saturating addition depending on the value of a special control register:

\begin{verbatim}
(parallel [(set (reg:SI 2) (unspec:SI [(reg:SI 3) (reg:SI 4)] 0)) (use (reg:SI 1))])
\end{verbatim}

This will not work, several of the optimizers only look at expressions locally; it is very likely that if you have multiple insns with identical inputs to the \texttt{unspec}, they will be optimized away even if register 1 changes in between.
This means that use can only be used to describe that the register is live. You should think twice before adding use statements, more often you will want to use espec instead. The use RTX is most commonly useful to describe that a fixed register is implicitly used in an insn. It is also safe to use in patterns where the compiler knows for other reasons that the result of the whole pattern is variable, such as ‘movmem’ or ‘call’ patterns.

During the reload phase, an insn that has a use as pattern can carry a reg_equal note. These use insns will be deleted before the reload phase exits.

During the delayed branch scheduling phase, x may be an insn. This indicates that x previously was located at this place in the code and its data dependencies need to be taken into account. These use insns will be deleted before the delayed branch scheduling phase exits.

\[
\text{(parallel [x0 x1 ...])}
\]

Represents several side effects performed in parallel. The square brackets stand for a vector; the operand of parallel is a vector of expressions. x0, x1 and so on are individual side effect expressions—expressions of code set, call, return, clobber or use.

“In parallel” means that first all the values used in the individual side-effects are computed, and second all the actual side-effects are performed. For example,

\[
\text{(parallel [(set (reg:SI 1) (mem:SI (reg:SI 1)))}
\text{(set (mem:SI (reg:SI 1))) (reg:SI 1)])}
\]

says unambiguously that the values of hard register 1 and the memory location addressed by it are interchanged. In both places where (reg:SI 1) appears as a memory address it refers to the value in register 1 before the execution of the insn.

It follows that it is incorrect to use parallel and expect the result of one set to be available for the next one. For example, people sometimes attempt to represent a jump-if-zero instruction this way:

\[
\text{(parallel [(set (cc0) (reg:SI 34))}
\text{(set (pc) (if_then_else}
\text{(eq (cc0) (const_int 0))}
\text{(label_ref ...) (pc)))])}
\]

But this is incorrect, because it says that the jump condition depends on the condition code value before this instruction, not on the new value that is set by this instruction.

Peephole optimization, which takes place together with final assembly code output, can produce insns whose patterns consist of a parallel whose elements are the operands needed to output the resulting assembler code—often reg, mem or constant expressions. This would not be well-formed RTL at any other stage in compilation, but it is ok then because no further optimization remains to be done. However, the definition of the macro NOTICE_UPDATE_CC, if any, must deal with such insns if you define any peephole optimizations.
(cond_exec [cond expr])
Represents a conditionally executed expression. The expr is executed only if the cond is nonzero. The cond expression must not have side-effects, but the expr may very well have side-effects.

(sequence [insns ...])
Represents a sequence of insns. Each of the insns that appears in the vector is suitable for appearing in the chain of insns, so it must be an insn, jump_insn, call_insn, code_label, barrier or note.
A sequence RTX is never placed in an actual insn during RTL generation. It represents the sequence of insns that result from a define_expand before those insns are passed to emit_insn to insert them in the chain of insns. When actually inserted, the individual sub-insns are separated out and the sequence is forgotten.
After delay-slot scheduling is completed, an insn and all the insns that reside in its delay slots are grouped together into a sequence. The insn requiring the delay slot is the first insn in the vector; subsequent insns are to be placed in the delay slot.
INSN_ANNULLED_BRANCH_P is set on an insn in a delay slot to indicate that a branch insn should be used that will conditionally annul the effect of the insns in the delay slots. In such a case, INSN_FROM_TARGET_P indicates that the insn is from the target of the branch and should be executed only if the branch is taken; otherwise the insn should be executed only if the branch is not taken.
See Section 16.19.7 [Delay Slots], page 379.

These expression codes appear in place of a side effect, as the body of an insn, though strictly speaking they do not always describe side effects as such:

(asm_input s)
Represents literal assembler code as described by the string s.

(unspec [operands ...] index)
(unspec_volatile [operands ...] index)
Represents a machine-specific operation on operands. index selects between multiple machine-specific operations. unspec_volatile is used for volatile operations and operations that may trap; unspec is used for other operations.
These codes may appear inside a pattern of an insn, inside a parallel, or inside an expression.

(addr_vec:m [lr0 lr1 ...])
Represents a table of jump addresses. The vector elements lr0, etc., are label_ref expressions. The mode m specifies how much space is given to each address; normally m would be Pmode.

(addr_diff_vec:m base [lr0 lr1 ...] min max flags)
Represents a table of jump addresses expressed as offsets from base. The vector elements lr0, etc., are label_ref expressions and so is base. The mode m specifies how much space is given to each address-difference. min and max are set up by branch shortening and hold a label with a minimum and a maximum
address, respectively. flags indicates the relative position of base, min and max to the containing insn and of min and max to base. See rtl.def for details.

(prefetch:m addr rw locality)
Represents prefetch of memory at address addr. Operand rw is 1 if the prefetch is for data to be written, 0 otherwise; targets that do not support write prefetches should treat this as a normal prefetch. Operand locality specifies the amount of temporal locality; 0 if there is none or 1, 2, or 3 for increasing levels of temporal locality; targets that do not support locality hints should ignore this.

This insn is used to minimize cache-miss latency by moving data into a cache before it is accessed. It should use only non-faulting data prefetch instructions.

10.16 Embedded Side-Effects on Addresses
Six special side-effect expression codes appear as memory addresses.

(pre_dec:m x)
Represents the side effect of decrementing x by a standard amount and represents also the value that x has after being decremented. x must be a reg or mem, but most machines allow only a reg. m must be the machine mode for pointers on the machine in use. The amount x is decremented by is the length in bytes of the machine mode of the containing memory reference of which this expression serves as the address. Here is an example of its use:

(mem:DF (pre_dec:SI (reg:SI 39)))
This says to decrement pseudo register 39 by the length of a DFmode value and use the result to address a DFmode value.

(pre_inc:m x)
Similar, but specifies incrementing x instead of decrementing it.

(post_dec:m x)
Represents the same side effect as pre_dec but a different value. The value represented here is the value x has before being decremented.

(post_inc:m x)
Similar, but specifies incrementing x instead of decrementing it.

(post_modify:m x y)
Represents the side effect of setting x to y and represents x before x is modified. x must be a reg or mem, but most machines allow only a reg. m must be the machine mode for pointers on the machine in use.

The expression y must be one of three forms: (plus:m x z), (minus:m x z), or (plus:m x i), where z is an index register and i is a constant.

Here is an example of its use:

(mem:SF (post_modify:SI (reg:SI 42) (plus (reg:SI 42)
                              (reg:SI 48))))
This says to modify pseudo register 42 by adding the contents of pseudo register 48 to it, after the use of what ever 42 points to.
(pre_modify:m x expr)

Similar except side effects happen before the use.

These embedded side effect expressions must be used with care. Instruction patterns may not use them. Until the ‘flow’ pass of the compiler, they may occur only to represent pushes onto the stack. The ‘flow’ pass finds cases where registers are incremented or decremented in one instruction and used as an address shortly before or after; these cases are then transformed to use pre- or post-increment or -decrement.

If a register used as the operand of these expressions is used in another address in an insn, the original value of the register is used. Uses of the register outside of an address are not permitted within the same insn as a use in an embedded side effect expression because such insns behave differently on different machines and hence must be treated as ambiguous and disallowed.

An instruction that can be represented with an embedded side effect could also be represented using parallel containing an additional set to describe how the address register is altered. This is not done because machines that allow these operations at all typically allow them wherever a memory address is called for. Describing them as additional parallel stores would require doubling the number of entries in the machine description.

10.17 Assembler Instructions as Expressions

The RTX code asm_operands represents a value produced by a user-specified assembler instruction. It is used to represent an asm statement with arguments. An asm statement with a single output operand, like this:

```asm
asm ("foo %1,%2,%0" : "=a" (outputvar) : "g" (x + y), "di" (*z));
```

is represented using a single asm_operands RTX which represents the value that is stored in outputvar:

```asm
(set rtx-for-outputvar
  (asm_operands "foo %1,%2,%0" "a" 0
   [rtx-for-addition-result rtx-for-*z]
   [(asm_input:mi "g")
    (asm_input:m2 "di")]))
```

Here the operands of the asm_operands RTX are the assembler template string, the output-operand’s constraint, the index-number of the output operand among the output operands specified, a vector of input operand RTX’s, and a vector of input-operand modes and constraints. The mode m1 is the mode of the sum x+y; m2 is that of *z.

When an asm statement has multiple output values, its insn has several such set RTX’s inside of a parallel. Each set contains an asm_operands; all of these share the same assembler template and vectors, but each contains the constraint for the respective output operand. They are also distinguished by the output-operand index number, which is 0, 1, ... for successive output operands.

10.18 Variable Location Debug Information in RTL

Variable tracking relies on MEM_EXPR and REG_EXPR annotations to determine what user variables memory and register references refer to.
Variable tracking at assignments uses these notes only when they refer to variables that live at fixed locations (e.g., addressable variables, global non-automatic variables). For variables whose location may vary, it relies on the following types of notes.

\[(\text{var\_location}:\text{mode} \ \text{var} \ \text{exp} \ \text{stat})\]

Binds variable \text{var}, a tree, to value \text{exp}, an RTL expression. It appears only in \text{NOTE\_INSN\_VAR\_LOCATION} and \text{DEBUG\_INSNs}, with slightly different meanings. \text{mode}, if present, represents the mode of \text{exp}, which is useful if it is a modeless expression. \text{stat} is only meaningful in notes, indicating whether the variable is known to be initialized or uninitialized.

\[(\text{debug\_expr}:\text{mode} \ \text{decl})\]

Stands for the value bound to the \text{DEBUG\_EXPR\_DECL} \text{decl}, that points back to it, within value expressions in \text{VAR\_LOCATION} nodes.

10.19 Insns

The RTL representation of the code for a function is a doubly-linked chain of objects called \text{insns}. Insns are expressions with special codes that are used for no other purpose. Some insns are actual instructions; others represent dispatch tables for \text{switch} statements; others represent labels to jump to or various sorts of declarative information.

In addition to its own specific data, each insn must have a unique id-number that distinguishes it from all other insns in the current function (after delayed branch scheduling, copies of an insn with the same id-number may be present in multiple places in a function, but these copies will always be identical and will only appear inside a \text{sequence}), and chain pointers to the preceding and following insns. These three fields occupy the same position in every insn, independent of the expression code of the insn. They could be accessed with \text{XEXP} and \text{XINT}, but instead three special macros are always used:

\text{INSN\_UID} (i)
Accesses the unique id of insn \text{i}.

\text{PREV\_INSN} (i)
Accesses the chain pointer to the insn preceding \text{i}. If \text{i} is the first insn, this is a null pointer.

\text{NEXT\_INSN} (i)
Accesses the chain pointer to the insn following \text{i}. If \text{i} is the last insn, this is a null pointer.

The first insn in the chain is obtained by calling \text{get\_insns}; the last insn is the result of calling \text{get\_last\_insn}. Within the chain delimited by these insns, the \text{NEXT\_INSN} and \text{PREV\_INSN} pointers must always correspond: if \text{insn} is not the first insn,

\[\text{NEXT\_INSN (PREV\_INSN (insn)) == insn}\]

is always true and if \text{insn} is not the last insn,

\[\text{PREV\_INSN (NEXT\_INSN (insn)) == insn}\]

is always true.

After delay slot scheduling, some of the insns in the chain might be \text{sequence} expressions, which contain a vector of insns. The value of \text{NEXT\_INSN} in all but the last of these insns is the next insn in the vector; the value of \text{NEXT\_INSN} of the last insn in the vector is the
same as the value of NEXT_INSN for the \texttt{sequence} in which it is contained. Similar rules apply for PREV_INSN.

This means that the above invariants are not necessarily true for insns inside \texttt{sequence} expressions. Specifically, if \texttt{insn} is the first insn in a \texttt{sequence}, NEXT_INSN (PREV_INSN (\texttt{insn})) is the insn containing the \texttt{sequence} expression, as is the value of PREV_INSN (NEXT_INSN (\texttt{insn})) if \texttt{insn} is the last insn in the \texttt{sequence} expression. You can use these expressions to find the containing \texttt{sequence} expression.

Every insn has one of the following expression codes:

\texttt{insn}  
\begin{itemize}
  \item The expression code \texttt{insn} is used for instructions that do not jump and do not do function calls. \texttt{sequence} expressions are always contained in insns with code \texttt{insn} even if one of those insns should jump or do function calls.
\end{itemize}

\texttt{insn} insns have four additional fields beyond the three mandatory ones listed above. These four are described in a table below.

\texttt{jump_insn}  
\begin{itemize}
  \item The expression code \texttt{jump_insn} is used for instructions that may jump (or, more generally, may contain label_ref expressions to which \texttt{pc} can be set in that instruction). If there is an instruction to return from the current function, it is recorded as a \texttt{jump_insn}.
\end{itemize}

\texttt{jump_insn} insns have the same extra fields as \texttt{insn} insns, accessed in the same way and in addition contain a field \texttt{JUMP_LABEL} which is defined once jump optimization has completed.

For simple conditional and unconditional jumps, this field contains the code\_label to which this insn will (possibly conditionally) branch. In a more complex jump, \texttt{JUMP_LABEL} records one of the labels that the insn refers to; other jump target labels are recorded as \texttt{REG_LABEL\_TARGET} notes. The exception is \texttt{addr\_vec} and \texttt{addr\_diff\_vec}, where \texttt{JUMP_LABEL} is \texttt{NULL\_RTX} and the only way to find the labels is to scan the entire body of the insn.

Return insns count as jumps, but since they do not refer to any labels, their \texttt{JUMP_LABEL} is \texttt{NULL\_RTX}.

\texttt{call_insn}  
\begin{itemize}
  \item The expression code \texttt{call_insn} is used for instructions that may do function calls. It is important to distinguish these instructions because they imply that certain registers and memory locations may be altered unpredictably.
\end{itemize}

\texttt{call_insn} insns have the same extra fields as \texttt{insn} insns, accessed in the same way and in addition contain a field \texttt{CALL\_INSN\_FUNCTION\_USAGE}, which contains a list (chain of \texttt{expr\_list} expressions) containing \texttt{use} and \texttt{clobber} expressions that denote hard registers and \texttt{MEM}s used or clobbered by the called function.

A \texttt{MEM} generally points to a stack slots in which arguments passed to the libcall by reference (see Section 17.10.7 [Register Arguments], page 449) are stored. If the argument is caller-copied (see Section 17.10.7 [Register Arguments], page 449), the stack slot will be mentioned in \texttt{Clobber} and \texttt{use} entries; if it’s callee-copied, only a \texttt{use} will appear, and the \texttt{MEM} may point to addresses that are not stack slots.
Clobbered registers in this list augment registers specified in CALL_USED_REGISTERS (see Section 17.7.1 [Register Basics], page 418).

code_label
A code_label insn represents a label that a jump insn can jump to. It contains two special fields of data in addition to the three standard ones. CODE_LABEL_NUMBER is used to hold the label number, a number that identifies this label uniquely among all the labels in the compilation (not just in the current function). Ultimately, the label is represented in the assembler output as an assembler label, usually of the form ‘Ln’ where n is the label number.

When a code_label appears in an RTL expression, it normally appears within a label_ref which represents the address of the label, as a number.

Besides as a code_label, a label can also be represented as a note of type NOTE_INSN_DELETED_LABEL.

The field LABEL_NUSES is only defined once the jump optimization phase is completed. It contains the number of times this label is referenced in the current function.

The field LABEL_KIND differentiates four different types of labels: LABEL_NORMAL, LABEL_STATIC_ENTRY, LABEL_GLOBAL_ENTRY, and LABEL_WEAK_ENTRY.

The only labels that do not have type LABEL_NORMAL are alternate entry points to the current function. These may be static (visible only in the containing translation unit), global (exposed to all translation units), or weak (global, but can be overridden by another symbol with the same name).

Much of the compiler treats all four kinds of label identically. Some of it needs to know whether or not a label is an alternate entry point; for this purpose, the macro LABEL_ALT_ENTRY_P is provided. It is equivalent to testing whether ‘LABEL_KIND (label) == LABEL_NORMAL’. The only place that cares about the distinction between static, global, and weak alternate entry points, besides the front-end code that creates them, is the function output_alternate_entry_point, in ‘final.c’.

To set the kind of a label, use the SET_LABEL_KIND macro.

barrier
Barriers are placed in the instruction stream when control cannot flow past them. They are placed after unconditional jump instructions to indicate that the jumps are unconditional and after calls to volatile functions, which do not return (e.g., exit). They contain no information beyond the three standard fields.

note
Note insns are used to represent additional debugging and declarative information. They contain two nonstandard fields, an integer which is accessed with the macro NOTE_LINE_NUMBER and a string accessed with NOTE_SOURCE_FILE.

If NOTE_LINE_NUMBER is positive, the note represents the position of a source line and NOTE_SOURCE_FILE is the source file name that the line came from. These notes control generation of line number data in the assembler output.

Otherwise, NOTE_LINE_NUMBER is not really a line number but a code with one of the following values (and NOTE_SOURCE_FILE must contain a null pointer):
NOTE_INSN_DELETED
Such a note is completely ignorable. Some passes of the compiler delete insns by altering them into notes of this kind.

NOTE_INSN_DELETED_LABEL
This marks what used to be a code_label, but was not used for other purposes than taking its address and was transformed to mark that no code jumps to it.

NOTE_INSN_BLOCK_BEG
NOTE_INSN_BLOCK_END
These types of notes indicate the position of the beginning and end of a level of scoping of variable names. They control the output of debugging information.

NOTE_INSN_EH_REGION_BEG
NOTE_INSN_EH_REGION_END
These types of notes indicate the position of the beginning and end of a level of scoping for exception handling. NOTE_BLOCK_NUMBER identifies which CODE_LABEL or note of type NOTE_INSN_DELETED_LABEL is associated with the given region.

NOTE_INSN_LOOP_BEG
NOTE_INSN_LOOP_END
These types of notes indicate the position of the beginning and end of a while or for loop. They enable the loop optimizer to find loops quickly.

NOTE_INSN_LOOP_CONT
Appears at the place in a loop that continue statements jump to.

NOTE_INSN_LOOP_VTOP
This note indicates the place in a loop where the exit test begins for those loops in which the exit test has been duplicated. This position becomes another virtual start of the loop when considering loop invariants.

NOTE_INSN_FUNCTION_BEG
Appears at the start of the function body, after the function prologue.

NOTE_INSN_VAR_LOCATION
This note is used to generate variable location debugging information. It indicates that the user variable in its VAR_LOCATION operand is at the location given in the RTL expression, or holds a value that can be computed by evaluating the RTL expression from that static point in the program up to the next such note for the same user variable.

These codes are printed symbolically when they appear in debugging dumps.
debug_insn

The expression code debug_insn is used for pseudo-instructions that hold debugging information for variable tracking at assignments (see ‘-fvar-tracking-assignments’ option). They are the RTL representation of GIMPLE_DEBUG statements (Section 12.7.7 [GIMPLE_DEBUG], page 239), with a VAR_LOCATION operand that binds a user variable tree to an RTL representation of the value in the corresponding statement. A DEBUG_EXPR in it stands for the value bound to the corresponding DEBUG_EXPR_DECL.

Throughout optimization passes, binding information is kept in pseudo-instruction form, so that, unlike notes, it gets the same treatment and adjustments that regular instructions would. It is the variable tracking pass that turns these pseudo-instructions into var location notes, analyzing control flow, value equivalences and changes to registers and memory referenced in value expressions, propagating the values of debug temporaries and determining expressions that can be used to compute the value of each user variable at as many points (ranges, actually) in the program as possible.

Unlike NOTE_INSN_VAR_LOCATION, the value expression in an INSN_VAR_LOCATION denotes a value at that specific point in the program, rather than an expression that can be evaluated at any later point before an overriding VAR_LOCATION is encountered. E.g., if a user variable is bound to a REG and then a subsequent insn modifies the REG, the note location would keep mapping the user variable to the register across the insn, whereas the insn location would keep the variable bound to the value, so that the variable tracking pass would emit another location note for the variable at the point in which the register is modified.

The machine mode of an insn is normally VOIDmode, but some phases use the mode for various purposes.

The common subexpression elimination pass sets the mode of an insn to QImode when it is the first insn in a block that has already been processed.

The second Haifa scheduling pass, for targets that can multiple issue, sets the mode of an insn to TImode when it is believed that the instruction begins an issue group. That is, when the instruction cannot issue simultaneously with the previous. This may be relied on by later passes, in particular machine-dependent reorg.

Here is a table of the extra fields of insn, jump_insn and call_insn insns:

**PATTERN (i)**
An expression for the side effect performed by this insn. This must be one of the following codes: set, call, use, clobber, return, asm_input, asm_output, addr_vec, addr_diff_vec, trap_if, unspec, unspec_volatile, parallel, cond_exec, or sequence. If it is a parallel, each element of the parallel must be one these codes, except that parallel expressions cannot be nested and addr_vec and addr_diff_vec are not permitted inside a parallel expression.

**INSN_CODE (i)**
An integer that says which pattern in the machine description matches this insn, or −1 if the matching has not yet been attempted.
Such matching is never attempted and this field remains −1 on an insn whose pattern consists of a single use, clobber, asm_input, addr_vec or addr_diff_vec expression.

Matching is also never attempted on insns that result from an asm statement. These contain at least one asm_operands expression. The function asm_noperands returns a non-negative value for such insns.

In the debugging output, this field is printed as a number followed by a symbolic representation that locates the pattern in the ‘md’ file as some small positive or negative offset from a named pattern.

LOG_LINKS (i)

A list (chain of insn_list expressions) giving information about dependencies between instructions within a basic block. Neither a jump nor a label may come between the related insns. These are only used by the schedulers and by combine. This is a deprecated data structure. Def-use and use-def chains are now preferred.

REG_NOTES (i)

A list (chain of expr_list and insn_list expressions) giving miscellaneous information about the insn. It is often information pertaining to the registers used in this insn.

The LOG_LINKS field of an insn is a chain of insn_list expressions. Each of these has two operands: the first is an insn, and the second is another insn_list expression (the next one in the chain). The last insn_list in the chain has a null pointer as second operand. The significant thing about the chain is which insns appear in it (as first operands of insn_list expressions). Their order is not significant.

This list is originally set up by the flow analysis pass; it is a null pointer until then. Flow only adds links for those data dependencies which can be used for instruction combination. For each insn, the flow analysis pass adds a link to insns which store into registers values that are used for the first time in this insn.

The REG_NOTES field of an insn is a chain similar to the LOG_LINKS field but it includes expr_list expressions in addition to insn_list expressions. There are several kinds of register notes, which are distinguished by the machine mode, which in a register note is really understood as being an enum reg_note. The first operand op of the note is data whose meaning depends on the kind of note.

The macro REG_NOTE_KIND (x) returns the kind of register note. Its counterpart, the macro PUT_REG_NOTE_KIND (x, newkind) sets the register note type of x to be newkind.

Register notes are of three classes: They may say something about an input to an insn, they may say something about an output of an insn, or they may create a linkage between two insns. There are also a set of values that are only used in LOG_LINKS.

These register notes annotate inputs to an insn:

REG_DEAD The value in op dies in this insn; that is to say, altering the value immediately after this insn would not affect the future behavior of the program.

It does not follow that the register op has no useful value after this insn since op is not necessarily modified by this insn. Rather, no subsequent instruction uses the contents of op.
The register \( op \) being set by this insn will not be used in a subsequent insn. This differs from a REG_DEAD note, which indicates that the value in an input will not be used subsequently. These two notes are independent; both may be present for the same register.

The register \( op \) is incremented (or decremented; at this level there is no distinction) by an embedded side effect inside this insn. This means it appears in a post_inc, pre_inc, post_dec or pre_dec expression.

The register \( op \) is known to have a nonnegative value when this insn is reached. This is used so that decrement and branch until zero instructions, such as the m68k dbra, can be matched.

The REG_NONNEG note is added to insns only if the machine description has a 'decrement_and_branch_until_zero' pattern.

This insn uses \( op \), a code_label or a note of type NOTE_INSN_DELETED_LABEL, but is not a jump_insn, or it is a jump_insn that refers to the operand as an ordinary operand. The label may still eventually be a jump target, but if so in an indirect jump in a subsequent insn. The presence of this note allows jump optimization to be aware that \( op \) is, in fact, being used, and flow optimization to build an accurate flow graph.

This insn is a jump_insn but not an addr_vec or addr_diff_vec. It uses \( op \), a code_label as a direct or indirect jump target. Its purpose is similar to that of REG_LABEL_OPERAND. This note is only present if the insn has multiple targets; the last label in the insn (in the highest numbered insn-field) goes into the JUMP_LABEL field and does not have a REG_LABEL_TARGET note. See Section 10.19 [Insns], page 166.

This insn is a branching instruction (either an unconditional jump or an indirect jump) which crosses between hot and cold sections, which could potentially be very far apart in the executable. The presence of this note indicates to other optimizations that this branching instruction should not be “collapsed” into a simpler branching construct. It is used when the optimization to partition basic blocks into hot and cold sections is turned on.

Appears attached to each CALL_INSN to setjmp or a related function.

The following notes describe attributes of outputs of an insn:

This note is only valid on an insn that sets only one register and indicates that that register will be equal to \( op \) at run time; the scope of this equivalence differs between the two types of notes. The value which the insn explicitly copies into
the register may look different from \textit{op}, but they will be equal at run time. If the output of the single \textit{set} is a \texttt{strict_low_part} expression, the note refers to the register that is contained in \texttt{SUBREG_REG} of the \texttt{subreg} expression.

For \texttt{REG_EQUIV}, the register is equivalent to \textit{op} throughout the entire function, and could validly be replaced in all its occurrences by \textit{op}. ("Validly" here refers to the data flow of the program; simple replacement may make some insns invalid.) For example, when a constant is loaded into a register that is never assigned any other value, this kind of note is used.

When a parameter is copied into a pseudo-register at entry to a function, a note of this kind records that the register is equivalent to the stack slot where the parameter was passed. Although in this case the register may be set by other insns, it is still valid to replace the register by the stack slot throughout the function.

A \texttt{REG_EQUIV} note is also used on an instruction which copies a register parameter into a pseudo-register at entry to a function, if there is a stack slot where that parameter could be stored. Although other insns may set the pseudo-register, it is valid for the compiler to replace the pseudo-register by stack slot throughout the function, provided the compiler ensures that the stack slot is properly initialized by making the replacement in the initial copy instruction as well. This is used on machines for which the calling convention allocates stack space for register parameters. See \texttt{REG_PARM_STACK_SPACE} in Section 17.10.6 [Stack Arguments], page 447.

In the case of \texttt{REG_EQUAL}, the register that is set by this insn will be equal to \textit{op} at run time at the end of this insn but not necessarily elsewhere in the function. In this case, \textit{op} is typically an arithmetic expression. For example, when a sequence of insns such as a library call is used to perform an arithmetic operation, this kind of note is attached to the insn that produces or copies the final value.

These two notes are used in different ways by the compiler passes. \texttt{REG_EQUAL} is used by passes prior to register allocation (such as common subexpression elimination and loop optimization) to tell them how to think of that value. \texttt{REG_EQUIV} notes are used by register allocation to indicate that there is an available substitute expression (either a constant or a \texttt{mem} expression for the location of a parameter on the stack) that may be used in place of a register if insufficient registers are available.

Except for stack homes for parameters, which are indicated by a \texttt{REG_EQUIV} note and are not useful to the early optimization passes and pseudo registers that are equivalent to a memory location throughout their entire life, which is not detected until later in the compilation, all equivalences are initially indicated by an attached \texttt{REG_EQUAL} note. In the early stages of register allocation, a \texttt{REG_EQUAL} note is changed into a \texttt{REG_EQUIV} note if \textit{op} is a constant and the insn represents the only set of its destination register.

Thus, compiler passes prior to register allocation need only check for \texttt{REG_EQUAL} notes and passes subsequent to register allocation need only check for \texttt{REG_EQUIV} notes.
These notes describe linkages between insns. They occur in pairs: one insn has one of a pair of notes that points to a second insn, which has the inverse note pointing back to the first insn.

**REG_CC_SETTER**
**REG_CC_USER**
On machines that use `cc0`, the insns which set and use `cc0` set and use `cc0` are adjacent. However, when branch delay slot filling is done, this may no longer be true. In this case a **REG_CC_USER** note will be placed on the insn setting `cc0` to point to the insn using `cc0` and a **REG_CC_SETTER** note will be placed on the insn using `cc0` to point to the insn setting `cc0`.

These values are only used in the **LOG_LINKS** field, and indicate the type of dependency that each link represents. Links which indicate a data dependence (a read after write dependence) do not use any code, they simply have mode `VOIDmode`, and are printed without any descriptive text.

**REG_DEP_TRUE**
This indicates a true dependence (a read after write dependence).

**REG_DEP_OUTPUT**
This indicates an output dependence (a write after write dependence).

**REG_DEP_ANTI**
This indicates an anti dependence (a write after read dependence).

These notes describe information gathered from gcov profile data. They are stored in the **REG_NOTES** field of an insn as an **expr_list**.

**REG_BR_PROB**
This is used to specify the ratio of branches to non-branches of a branch insn according to the profile data. The value is stored as a value between 0 and **REG_BR_PROB_BASE**; larger values indicate a higher probability that the branch will be taken.

**REG_BR_PRED**
These notes are found in JUMP insns after delayed branch scheduling has taken place. They indicate both the direction and the likelihood of the JUMP. The format is a bitmask of **ATTR_FLAG** values.

**REG_FRAME_RELATED_EXPR**
This is used on an RTX_FRAME_RELATED_P insn wherein the attached expression is used in place of the actual insn pattern. This is done in cases where the pattern is either complex or misleading.

For convenience, the machine mode in an **insn_list** or **expr_list** is printed using these symbolic codes in debugging dumps.

The only difference between the expression codes **insn_list** and **expr_list** is that the first operand of an **insn_list** is assumed to be an insn and is printed in debugging dumps as the insn's unique id; the first operand of an **expr_list** is printed in the ordinary way as an expression.
10.20 RTL Representation of Function-Call Insns

Insns that call subroutines have the RTL expression code `call_insn`. These insns must satisfy special rules, and their bodies must use a special RTL expression code, `call`.

A `call` expression has two operands, as follows:

```
(call (mem:fm addr) nbytes)
```

Here `nbytes` is an operand that represents the number of bytes of argument data being passed to the subroutine, `fm` is a machine mode (which must equal as the definition of the `FUNCTION_MODE` macro in the machine description) and `addr` represents the address of the subroutine.

For a subroutine that returns no value, the `call` expression as shown above is the entire body of the insn, except that the insn might also contain `use` or `clobber` expressions.

For a subroutine that returns a value whose mode is not `BLKmode`, the value is returned in a hard register. If this register’s number is `r`, then the body of the call insn looks like this:

```
(set (reg:m r)
     (call (mem:fm addr) nbytes))
```

This RTL expression makes it clear (to the optimizer passes) that the appropriate register receives a useful value in this insn.

When a subroutine returns a `BLKmode` value, it is handled by passing to the subroutine the address of a place to store the value. So the call insn itself does not “return” any value, and it has the same RTL form as a call that returns nothing.

On some machines, the call instruction itself clobbers some register, for example to contain the return address. `call_insn` insns on these machines should have a body which is a `parallel` that contains both the `call` expression and `clobber` expressions that indicate which registers are destroyed. Similarly, if the call instruction requires some register other than the stack pointer that is not explicitly mentioned in its RTL, a `use` subexpression should mention that register.

Functions that are called are assumed to modify all registers listed in the configuration macro `CALL_USED_REGISTERS` (see Section 17.7.1 [Register Basics], page 418) and, with the exception of `const` functions and library calls, to modify all of memory.

Insns containing just `use` expressions directly precede the `call_insn` insn to indicate which registers contain inputs to the function. Similarly, if registers other than those in `CALL_USED_REGISTERS` are clobbered by the called function, insns containing a single `clobber` follow immediately after the call to indicate which registers.

10.21 Structure Sharing Assumptions

The compiler assumes that certain kinds of RTL expressions are unique; there do not exist two distinct objects representing the same value. In other cases, it makes an opposite assumption: that no RTL expression object of a certain kind appears in more than one place in the containing structure.

These assumptions refer to a single function; except for the RTL objects that describe global variables and external functions, and a few standard objects such as small integer constants, no RTL objects are common to two functions.
Each pseudo-register has only a single \texttt{reg} object to represent it, and therefore only a single machine mode.

- For any symbolic label, there is only one \texttt{symbol_ref} object referring to it.
- All \texttt{const_int} expressions with equal values are shared.
- There is only one \texttt{pc} expression.
- There is only one \texttt{cc0} expression.
- There is only one \texttt{const_double} expression with value 0 for each floating point mode. Likewise for values 1 and 2.
- There is only one \texttt{const_vector} expression with value 0 for each vector mode, be it an integer or a double constant vector.
- No \texttt{label_ref} or \texttt{scratch} appears in more than one place in the RTL structure; in other words, it is safe to do a tree-walk of all the insns in the function and assume that each time a \texttt{label_ref} or \texttt{scratch} is seen it is distinct from all others that are seen.
- Only one \texttt{mem} object is normally created for each static variable or stack slot, so these objects are frequently shared in all the places they appear. However, separate but equal objects for these variables are occasionally made.
- When a single \texttt{asm} statement has multiple output operands, a distinct \texttt{asm_operands} expression is made for each output operand. However, these all share the vector which contains the sequence of input operands. This sharing is used later on to test whether two \texttt{asm_operands} expressions come from the same statement, so all optimizations must carefully preserve the sharing if they copy the vector at all.
- No RTL object appears in more than one place in the RTL structure except as described above. Many passes of the compiler rely on this by assuming that they can modify RTL objects in place without unwanted side-effects on other insns.
- During initial RTL generation, shared structure is freely introduced. After all the RTL for a function has been generated, all shared structure is copied by \texttt{unshare_all_rtl} in \texttt{emit-rtl.c}, after which the above rules are guaranteed to be followed.
- During the combiner pass, shared structure within an insn can exist temporarily. However, the shared structure is copied before the combiner is finished with the insn. This is done by calling \texttt{copy_rtx_if_shared}, which is a subroutine of \texttt{unshare_all_rtl}.

### 10.22 Reading RTL

To read an RTL object from a file, call \texttt{read_rtx}. It takes one argument, a stdio stream, and returns a single RTL object. This routine is defined in \texttt{read-rtl.c}. It is not available in the compiler itself, only the various programs that generate the compiler back end from the machine description.

People frequently have the idea of using RTL stored as text in a file as an interface between a language front end and the bulk of GCC. This idea is not feasible.

GCC was designed to use RTL internally only. Correct RTL for a given program is very dependent on the particular target machine. And the RTL does not contain all the information about the program.

The proper way to interface GCC to a new language front end is with the “tree” data structure, described in the files \texttt{tree.h} and \texttt{tree.def}. The documentation for this structure (see Chapter 11 [GENERIC], page 177) is incomplete.
11 GENERIC

The purpose of GENERIC is simply to provide a language-independent way of representing an entire function in trees. To this end, it was necessary to add a few new tree codes to the back end, but most everything was already there. If you can express it with the codes in gcc/tree.def, it’s GENERIC.

Early on, there was a great deal of debate about how to think about statements in a tree IL. In GENERIC, a statement is defined as any expression whose value, if any, is ignored. A statement will always have TREE_SIDE_EFFECTS set (or it will be discarded), but a non-statement expression may also have side effects. A CALL_EXPR, for instance.

It would be possible for some local optimizations to work on the GENERIC form of a function; indeed, the adapted tree inliner works fine on GENERIC, but the current compiler performs inlining after lowering to GIMPLE (a restricted form described in the next section). Indeed, currently the frontends perform this lowering before handing off to tree_rest_of_compilation, but this seems inelegant.

11.1 Deficiencies

There are many places in which this document is incomplet and incorrekt. It is, as of yet, only preliminary documentation.

11.2 Overview

The central data structure used by the internal representation is the tree. These nodes, while all of the C type tree, are of many varieties. A tree is a pointer type, but the object to which it points may be of a variety of types. From this point forward, we will refer to trees in ordinary type, rather than in this font, except when talking about the actual C type tree.

You can tell what kind of node a particular tree is by using the TREE_CODE macro. Many, many macros take trees as input and return trees as output. However, most macros require a certain kind of tree node as input. In other words, there is a type-system for trees, but it is not reflected in the C type-system.

For safety, it is useful to configure GCC with ‘--enable-checking’. Although this results in a significant performance penalty (since all tree types are checked at run-time), and is therefore inappropriate in a release version, it is extremely helpful during the development process.

Many macros behave as predicates. Many, although not all, of these predicates end in ‘_P’. Do not rely on the result type of these macros being of any particular type. You may, however, rely on the fact that the type can be compared to 0, so that statements like

\[
\text{if (TEST_P (t) && !TEST_P (y))}
\]

\[
x = 1;
\]

and

\[
\text{int i = (TEST_P (t) != 0);}
\]

are legal. Macros that return int values now may be changed to return tree values, or other pointers in the future. Even those that continue to return int may return multiple nonzero codes where previously they returned only zero and one. Therefore, you should not write code like
if (TEST_P (t) == 1)
as this code is not guaranteed to work correctly in the future.

You should not take the address of values returned by the macros or functions described here. In particular, no guarantee is given that the values are lvalues.

In general, the names of macros are all in uppercase, while the names of functions are entirely in lowercase. There are rare exceptions to this rule. You should assume that any macro or function whose name is made up entirely of uppercase letters may evaluate its arguments more than once. You may assume that a macro or function whose name is made up entirely of lowercase letters will evaluate its arguments only once.

The error_mark_node is a special tree. Its tree code is ERROR_MARK, but since there is only ever one node with that code, the usual practice is to compare the tree against error_mark_node. (This test is just a test for pointer equality.) If an error has occurred during front-end processing the flag errorcount will be set. If the front end has encountered code it cannot handle, it will issue a message to the user and set sorrycount. When these flags are set, any macro or function which normally returns a tree of a particular kind may instead return the error_mark_node. Thus, if you intend to do any processing of erroneous code, you must be prepared to deal with the error_mark_node.

Occasionally, a particular tree slot (like an operand to an expression, or a particular field in a declaration) will be referred to as “reserved for the back end”. These slots are used to store RTL when the tree is converted to RTL for use by the GCC back end. However, if that process is not taking place (e.g., if the front end is being hooked up to an intelligent editor), then those slots may be used by the back end presently in use.

If you encounter situations that do not match this documentation, such as tree nodes of types not mentioned here, or macros documented to return entities of a particular kind that instead return entities of some different kind, you have found a bug, either in the front end or in the documentation. Please report these bugs as you would any other bug.

11.2.1 Trees
All GENERIC trees have two fields in common. First, TREE_CHAIN is a pointer that can be used as a singly-linked list to other trees. The other is TREE_TYPE. Many trees store the type of an expression or declaration in this field.

These are some other functions for handling trees:

```
tree_size
```

> Return the number of bytes a tree takes.

build0
build1
build2
build3
build4
build5
build6

> These functions build a tree and supply values to put in each parameter. The basic signature is `code, type, [operands]`. code is the TREE_CODE, and type is a tree representing the TREE_TYPE. These are followed by the operands, each of which is also a tree.
11.2.2 Identifiers

An IDENTIFIER_NODE represents a slightly more general concept that the standard C or C++ concept of identifier. In particular, an IDENTIFIER_NODE may contain a ‘$’, or other extraordinary characters.

There are never two distinct IDENTIFIER_NODEs representing the same identifier. Therefore, you may use pointer equality to compare IDENTIFIER_NODEs, rather than using a routine like strcmp. Use get_identifier to obtain the unique IDENTIFIER_NODE for a supplied string.

You can use the following macros to access identifiers:

IDENTIFIER_POINTER
The string represented by the identifier, represented as a char*. This string is always NUL-terminated, and contains no embedded NUL characters.

IDENTIFIER_LENGTH
The length of the string returned by IDENTIFIER_POINTER, not including the trailing NUL. This value of IDENTIFIER_LENGTH (x) is always the same as strlen (IDENTIFIER_POINTER (x)).

IDENTIFIER_OPNAME_P
This predicate holds if the identifier represents the name of an overloaded operator. In this case, you should not depend on the contents of either the IDENTIFIER_POINTER or the IDENTIFIER_LENGTH.

IDENTIFIER_TYPENAME_P
This predicate holds if the identifier represents the name of a user-defined conversion operator. In this case, the TREE_TYPE of the IDENTIFIER_NODE holds the type to which the conversion operator converts.

11.2.3 Containers

Two common container data structures can be represented directly with tree nodes. A TREE_LIST is a singly linked list containing two trees per node. These are the TREE_PURPOSE and TREE_VALUE of each node. (Often, the TREE_PURPOSE contains some kind of tag, or additional information, while the TREE_VALUE contains the majority of the payload. In other cases, the TREE_PURPOSE is simply NULL_TREE, while in still others both the TREE_PURPOSE and TREE_VALUE are of equal stature.) Given one TREE_LIST node, the next node is found by following the TREE_CHAIN. If the TREE_CHAIN is NULL_TREE, then you have reached the end of the list.

A TREE_VEC is a simple vector. The TREE_VEC_LENGTH is an integer (not a tree) giving the number of nodes in the vector. The nodes themselves are accessed using the TREE_VEC_ELT macro, which takes two arguments. The first is the TREE_VEC in question; the second is an integer indicating which element in the vector is desired. The elements are indexed from zero.

11.3 Types

All types have corresponding tree nodes. However, you should not assume that there is exactly one tree node corresponding to each type. There are often multiple nodes corresponding to the same type.
For the most part, different kinds of types have different tree codes. (For example, pointer types use a **POINTER_TYPE** code while arrays use an **ARRAY_TYPE** code.) However, pointers to member functions use the **RECORD_TYPE** code. Therefore, when writing a **switch** statement that depends on the code associated with a particular type, you should take care to handle pointers to member functions under the **RECORD_TYPE** case label.

The following functions and macros deal with cv-qualification of types:

**TYPE_MAIN_VARIANT**
This macro returns the unqualified version of a type. It may be applied to an unqualified type, but it is not always the identity function in that case.

A few other macros and functions are usable with all types:

**TYPE_SIZE**
The number of bits required to represent the type, represented as an **INTEGER_CST**. For an incomplete type, **TYPE_SIZE** will be **NULL_TREE**.

**TYPE_ALIGN**
The alignment of the type, in bits, represented as an **int**.

**TYPE_NAME**
This macro returns a declaration (in the form of a **TYPE_DECL**) for the type. (Note this macro does **not** return an **IDENTIFIER_NODE**, as you might expect, given its name!) You can look at the **DECL_NAME** of the **TYPE_DECL** to obtain the actual name of the type. The **TYPE_NAME** will be **NULL_TREE** for a type that is not a built-in type, the result of a typedef, or a named class type.

**TYPE_CANONICAL**
This macro returns the “canonical” type for the given type node. Canonical types are used to improve performance in the C++ and Objective-C++ front ends by allowing efficient comparison between two type nodes in **same_type_p**: if the **TYPE_CANONICAL** values of the types are equal, the types are equivalent; otherwise, the types are not equivalent. The notion of equivalence for canonical types is the same as the notion of type equivalence in the language itself. For instance,

When **TYPE_CANONICAL** is **NULL_TREE**, there is no canonical type for the given type node. In this case, comparison between this type and any other type requires the compiler to perform a deep, “structural” comparison to see if the two type nodes have the same form and properties.

The canonical type for a node is always the most fundamental type in the equivalence class of types. For instance, **int** is its own canonical type. A typedef **I** of **int** will have **int** as its canonical type. Similarly, **I** and a typedef **IP** (defined to **I** *) will have **int** * as their canonical type. When building a new type node, be sure to set **TYPE_CANONICAL** to the appropriate canonical type.

If the new type is a compound type (built from other types), and any of those other types require structural equality, use **SET_TYPE_STRUCTURAL_EQUALITY** to ensure that the new type also requires structural equality. Finally, if for some reason you cannot guarantee that **TYPE_CANONICAL** will point to the canonical type, use **SET_TYPE_STRUCTURAL_EQUALITY** to make sure that the new type—and any type constructed based on it—requires structural equality. If you suspect
that the canonical type system is miscomparing types, pass `--param verify-canonical-types=1` to the compiler or configure with `--enable-checking` to force the compiler to verify its canonical-type comparisons against the structural comparisons; the compiler will then print any warnings if the canonical types miscompare.

**TYPE_STRUCTURAL_EQUALITY_P**
This predicate holds when the node requires structural equality checks, e.g., when TYPE_CANONICAL is NULL_TREE.

**SET_TYPE_STRUCTURAL_EQUALITY**
This macro states that the type node it is given requires structural equality checks, e.g., it sets TYPE_CANONICAL to NULL_TREE.

**same_type_p**
This predicate takes two types as input, and holds if they are the same type. For example, if one type is a typedef for the other, or both are typedefs for the same type. This predicate also holds if the two trees given as input are simply copies of one another; i.e., there is no difference between them at the source level, but, for whatever reason, a duplicate has been made in the representation. You should never use `==` (pointer equality) to compare types; always use `same_type_p` instead.

Detailed below are the various kinds of types, and the macros that can be used to access them. Although other kinds of types are used elsewhere in G++, the types described here are the only ones that you will encounter while examining the intermediate representation.

**VOID_TYPE**
Used to represent the void type.

**INTEGER_TYPE**
Used to represent the various integral types, including char, short, int, long, and long long. This code is not used for enumeration types, nor for the bool type. The TYPE_PRECISION is the number of bits used in the representation, represented as an unsigned int. (Note that in the general case this is not the same value as TYPE_SIZE; suppose that there were a 24-bit integer type, but that alignment requirements for the ABI required 32-bit alignment. Then, TYPE_SIZE would be an INTEGER_CST for 32, while TYPE_PRECISION would be 24.) The integer type is unsigned if TYPE_UNSIGNED holds; otherwise, it is signed.

The TYPE_MIN_VALUE is an INTEGER_CST for the smallest integer that may be represented by this type. Similarly, the TYPE_MAX_VALUE is an INTEGER_CST for the largest integer that may be represented by this type.

**REAL_TYPE**
Used to represent the float, double, and long double types. The number of bits in the floating-point representation is given by TYPE_PRECISION, as in the INTEGER_TYPE case.

**FIXED_POINT_TYPE**
Used to represent the short _Fract, _Fract, long _Fract, long long _Fract, short _Accum, _Accum, long _Accum, and long long _Accum types. The num-
ber of bits in the fixed-point representation is given by `TYPE_PRECISION`, as in the `INTEGER_TYPE` case. There may be padding bits, fractional bits and integral bits. The number of fractional bits is given by `TYPE_FBIT`, and the number of integral bits is given by `TYPE_IBIT`. The fixed-point type is unsigned if `TYPE_UNSIGNED` holds; otherwise, it is signed. The fixed-point type is saturating if `TYPE_SATURATING` holds; otherwise, it is not saturating.

**COMPLEX_TYPE**

Used to represent GCC built-in `__complex__` data types. The `TREE_TYPE` is the type of the real and imaginary parts.

**ENUMERAL_TYPE**

Used to represent an enumeration type. The `TYPE_PRECISION` gives (as an `int`), the number of bits used to represent the type. If there are no negative enumeration constants, `TYPE_UNSIGNED` will hold. The minimum and maximum enumeration constants may be obtained with `TYPE_MIN_VALUE` and `TYPE_MAX_VALUE`, respectively; each of these macros returns an `INTEGER_CST`.

The actual enumeration constants themselves may be obtained by looking at the `TYPE_VALUES`. This macro will return a `TREE_LIST`, containing the constants. The `TREE_PURPOSE` of each node will be an `IDENTIFIER_NODE` giving the name of the constant; the `TREE_VALUE` will be an `INTEGER_CST` giving the value assigned to that constant. These constants will appear in the order in which they were declared. The `TREE_TYPE` of each of these constants will be the type of enumeration type itself.

**BOOLEAN_TYPE**

Used to represent the `bool` type.

**POINTER_TYPE**

Used to represent pointer types, and pointer to data member types. The `TREE_TYPE` gives the type to which this type points.

**REFERENCE_TYPE**

Used to represent reference types. The `TREE_TYPE` gives the type to which this type refers.

**FUNCTION_TYPE**

Used to represent the type of non-member functions and of static member functions. The `TREE_TYPE` gives the return type of the function. The `TYPE_ARG_TYPES` are a `TREE_LIST` of the argument types. The `TREE_VALUE` of each node in this list is the type of the corresponding argument; the `TREE_PURPOSE` is an expression for the default argument value, if any. If the last node in the list is `void_list_node` (a `TREE_LIST` node whose `TREE_VALUE` is the `void_type_node`), then functions of this type do not take variable arguments. Otherwise, they do take a variable number of arguments.

Note that in C (but not in C++) a function declared like `void f()` is an unprototyped function taking a variable number of arguments; the `TYPE_ARG_TYPES` of such a function will be `NULL`. 
METHOD_TYPE
Used to represent the type of a non-static member function. Like a FUNCTION_TYPE, the return type is given by the TREE_TYPE. The type of *this, i.e., the class of which functions of this type are a member, is given by the TYPE_METHOD_BASETYPE. The TYPE_ARG_TYPES is the parameter list, as for a FUNCTION_TYPE, and includes the this argument.

ARRAY_TYPE
Used to represent array types. The TREE_TYPE gives the type of the elements in the array. If the array-bound is present in the type, the TYPE_DOMAIN is an INTEGER_TYPE whose TYPE_MIN_VALUE and TYPE_MAX_VALUE will be the lower and upper bounds of the array, respectively. The TYPE_MIN_VALUE will always be an INTEGER_CST for zero, while the TYPE_MAX_VALUE will be one less than the number of elements in the array, i.e., the highest value which may be used to index an element in the array.

RECORD_TYPE
Used to represent struct and class types, as well as pointers to member functions and similar constructs in other languages. TYPE_FIELDS contains the items contained in this type, each of which can be a FIELD_DECL, VAR_DECL, CONST_DECL, or TYPE_DECL. You may not make any assumptions about the ordering of the fields in the type or whether one or more of them overlap.

UNION_TYPE
Used to represent union types. Similar to RECORD_TYPE except that all FIELD_DECL nodes in TYPE_FIELD start at bit position zero.

QUAL_UNION_TYPE
Used to represent part of a variant record in Ada. Similar to UNION_TYPE except that each FIELD_DECL has a DECL_QUALIFIER field, which contains a boolean expression that indicates whether the field is present in the object. The type will only have one field, so each field’s DECL_QUALIFIER is only evaluated if none of the expressions in the previous fields in TYPE_FIELDS are nonzero. Normally these expressions will reference a field in the outer object using a PLACEHOLDER_EXPR.

LANG_TYPE
This node is used to represent a language-specific type. The front end must handle it.

OFFSET_TYPE
This node is used to represent a pointer-to-data member. For a data member X::m the TYPE_OFFSET_BASETYPE is X and the TREE_TYPE is the type of m.

There are variables whose values represent some of the basic types. These include:

void_type_node
A node for void.

integer_type_node
A node for int.
unsigned_type_node.
   A node for unsigned int.

char_type_node.
   A node for char.

It may sometimes be useful to compare one of these variables with a type in hand, using same_type_p.

11.4 Declarations

This section covers the various kinds of declarations that appear in the internal representation, except for declarations of functions (represented by FUNCTION_DECL nodes), which are described in Section 11.8 [Functions], page 206.

11.4.1 Working with declarations

Some macros can be used with any kind of declaration. These include:

DECL_NAME
   This macro returns an IDENTIFIER_NODE giving the name of the entity.

TREE_TYPE
   This macro returns the type of the entity declared.

EXPR_FILENAME
   This macro returns the name of the file in which the entity was declared, as a char*. For an entity declared implicitly by the compiler (like __builtin_memcpy), this will be the string "<internal>".

EXPR_LINENO
   This macro returns the line number at which the entity was declared, as an int.

DECL_ARTIFICIAL
   This predicate holds if the declaration was implicitly generated by the compiler. For example, this predicate will hold of an implicitly declared member function, or of the TYPEDECL implicitly generated for a class type. Recall that in C++ code like:
      
      struct S {};
      
      is roughly equivalent to C code like:
      
      struct S {};
      typedef struct S S;

      The implicitly generated typedef declaration is represented by a TYPEDECL for which DECL_ARTIFICIAL holds.

   The various kinds of declarations include:

LABELDECL
   These nodes are used to represent labels in function bodies. For more information, see Section 11.8 [Functions], page 206. These nodes only appear in block scopes.
CONST_DECL
These nodes are used to represent enumeration constants. The value of the constant is given by *DECL_INITIAL* which will be an *INTEGER_CST* with the same type as the *TREE_TYPE* of the *CONST_DECL*, i.e., an *ENUMERAL_TYPE*.

RESULT_DECL
These nodes represent the value returned by a function. When a value is assigned to a *RESULT_DECL*, that indicates that the value should be returned, via bitwise copy, by the function. You can use *DECL_SIZE* and *DECL_ALIGN* on a *RESULT_DECL*, just as with a *VAR_DECL*.

TYPE_DECL
These nodes represent *typedef* declarations. The *TREE_TYPE* is the type declared to have the name given by *DECL_NAME*. In some cases, there is no associated name.

VAR_DECL
These nodes represent variables with namespace or block scope, as well as static data members. The *DECL_SIZE* and *DECL_ALIGN* are analogous to *TYPE_SIZE* and *TYPE_ALIGN*. For a declaration, you should always use the *DECL_SIZE* and *DECL_ALIGN* rather than the *TYPE_SIZE* and *TYPE_ALIGN* given by the *TREE_TYPE*, since special attributes may have been applied to the variable to give it a particular size and alignment. You may use the predicates *DECL_THIS_STATIC* or *DECL_THIS_EXTERN* to test whether the storage class specifiers *static* or *extern* were used to declare a variable.

If this variable is initialized (but does not require a constructor), the *DECL_INITIAL* will be an expression for the initializer. The initializer should be evaluated, and a bitwise copy into the variable performed. If the *DECL_INITIAL* is the *error_mark_node*, there is an initializer, but it is given by an explicit statement later in the code; no bitwise copy is required.

GCC provides an extension that allows either automatic variables, or global variables, to be placed in particular registers. This extension is being used for a particular *VAR_DECL* if *DECL_REGISTER* holds for the *VAR_DECL*, and if *DECL_ASSEMBLER_NAME* is not equal to *DECL_NAME*. In that case, *DECL_ASSEMBLER_NAME* is the name of the register into which the variable will be placed.

PARM_DECL
Used to represent a parameter to a function. Treat these nodes similarly to *VAR_DECL* nodes. These nodes only appear in the *DECL_ARGUMENTS* for a *FUNCTION_DECL*.

The *DECL_ARG_TYPE* for a *PARM_DECL* is the type that will actually be used when a value is passed to this function. It may be a wider type than the *TREE_TYPE* of the parameter; for example, the ordinary type might be *short* while the *DECL_ARG_TYPE* is *int*.

DEBUG_EXPR_DECL
Used to represent an anonymous debug-information temporary created to hold an expression as it is optimized away, so that its value can be referenced in debug bind statements.
FIELD_DECL
These nodes represent non-static data members. The DECL_SIZE and DECL_ALIGN behave as for VAR_DECL nodes. The position of the field within the parent record is specified by a combination of three attributes. DECL_FIELD_OFFSET is the position, counting in bytes, of the DECL_OFFSET_ALIGN-bit sized word containing the bit of the field closest to the beginning of the structure. DECL_FIELD_BIT_OFFSET is the bit offset of the first bit of the field within this word; this may be nonzero even for fields that are not bit-fields, since DECL_OFFSET_ALIGN may be greater than the natural alignment of the field’s type.
If DECL_C_BIT_FIELD holds, this field is a bit-field. In a bit-field, DECL_BIT_FIELD_TYPE also contains the type that was originally specified for it, while DECL_TYPE may be a modified type with lesser precision, according to the size of the bit field.

NAMESPACE_DECL
Namespaces provide a name hierarchy for other declarations. They appear in the DECL_CONTEXT of other DECL nodes.

11.4.2 Internal structure
DECL nodes are represented internally as a hierarchy of structures.

11.4.2.1 Current structure hierarchy
struct tree_decl_minimal
This is the minimal structure to inherit from in order for common DECL macros to work. The fields it contains are a unique ID, source location, context, and name.

struct tree_decl_common
This structure inherits from struct tree_decl_minimal. It contains fields that most DECL nodes need, such as a field to store alignment, machine mode, size, and attributes.

struct tree_field_decl
This structure inherits from struct tree_decl_common. It is used to represent FIELD_DECL.

struct tree_label_decl
This structure inherits from struct tree_decl_common. It is used to represent LABEL_DECL.

struct tree_translation_unit_decl
This structure inherits from struct tree_decl_common. It is used to represent TRANSLATION_UNIT_DECL.

struct tree_decl_with_rtl
This structure inherits from struct tree_decl_common. It contains a field to store the low-level RTL associated with a DECL node.

struct tree_result_decl
This structure inherits from struct tree_decl_with_rtl. It is used to represent RESULT_DECL.
struct tree_const_decl
This structure inherits from struct tree_decl_with_rtl. It is used to represent CONST_DECL.

struct tree_parm_decl
This structure inherits from struct tree_decl_with_rtl. It is used to represent PARM_DECL.

struct tree_decl_with_vis
This structure inherits from struct tree_decl_with_rtl. It contains fields necessary to store visibility information, as well as a section name and assembler name.

struct tree_var_decl
This structure inherits from struct tree_decl_with_vis. It is used to represent VAR_DECL.

struct tree_function_decl
This structure inherits from struct tree_decl_with_vis. It is used to represent FUNCTION_DECL.

11.4.2.2 Adding new DECL node types
Adding a new DECL tree consists of the following steps

Add a new tree code for the DECL node
For language specific DECL nodes, there is a `.def` file in each frontend directory where the tree code should be added. For DECL nodes that are part of the middle-end, the code should be added to `tree.def`.

Create a new structure type for the DECL node
These structures should inherit from one of the existing structures in the language hierarchy by using that structure as the first member.

```c
struct tree_foo_decl
{
    struct tree_decl_with_vis common;
}
```

Would create a structure name `tree_foo_decl` that inherits from `struct tree_decl_with_vis`.

For language specific DECL nodes, this new structure type should go in the appropriate `.h` file. For DECL nodes that are part of the middle-end, the structure type should go in `tree.h`.

Add a member to the tree structure enumerator for the node
For garbage collection and dynamic checking purposes, each DECL node structure type is required to have a unique enumerator value specified with it. For language specific DECL nodes, this new enumerator value should go in the appropriate `.def` file. For DECL nodes that are part of the middle-end, the enumerator values are specified in `treestruct.def`. 
Update **union tree_node**

In order to make your new structure type usable, it must be added to **union tree_node**. For language specific DECL nodes, a new entry should be added to the appropriate `.h` file of the form

```c
struct tree_foo_decl GTY ((tag ("TS_VAR_DECL"))) foo_decl;
```

For DECL nodes that are part of the middle-end, the additional member goes directly into **union tree_node** in `tree.h`.

Update dynamic checking info

In order to be able to check whether accessing a named portion of **union tree_node** is legal, and whether a certain DECL node contains one of the enumerated DECL node structures in the hierarchy, a simple lookup table is used. This lookup table needs to be kept up to date with the tree structure hierarchy, or else checking and containment macros will fail inappropriately.

For language specific DECL nodes, there is an `init_ts` function in an appropriate `.c` file, which initializes the lookup table. Code setting up the table for new DECL nodes should be added there. For each DECL tree code and enumerator value representing a member of the inheritance hierarchy, the table should contain 1 if that tree code inherits (directly or indirectly) from that member. Thus, a FOO_DECL node derived from **struct decl_with rtl**, and enumerator value TS_FOO_DECL, would be set up as follows

```c
tree_contains_struct[FOO_DECL][TS_FOO_DECL] = 1;
tree_contains_struct[FOO_DECL][TS_DECL_WRTL] = 1;
```

For DECL nodes that are part of the middle-end, the setup code goes into `tree.c`.

Add macros to access any new fields and flags

Each added field or flag should have a macro that is used to access it, that performs appropriate checking to ensure only the right type of DECL nodes access the field.

These macros generally take the following form

```c
#define FOO_DECL_FIELDNAME(NODE) FOO_DECL_CHECK(NODE)->foo_decl.fieldname
```

However, if the structure is simply a base class for further structures, something like the following should be used

```c
#define BASE_STRUCT_CHECK(T) CONTAINS_STRUCT_CHECK(T, TS_BASE_STRUCT)
#define BASE_STRUCT_FIELDNAME(NODE) \  (BASE_STRUCT_CHECK(NODE)->base_struct.fieldname
```

### 11.5 Attributes in trees

Attributes, as specified using the `__attribute__` keyword, are represented internally as a **TREE_LIST**. The **TREE_PURPOSE** is the name of the attribute, as an **IDENTIFIER_NODE**. The **TREE_VALUE** is a **TREE_LIST** of the arguments of the attribute, if any, or **NULL_TREE** if there are no arguments; the arguments are stored as the **TREE_VALUE** of successive entries in the list, and may be identifiers or expressions. The **TREE_CHAIN** of the attribute is the next attribute in a list of attributes applying to the same declaration or type, or **NULL_TREE** if there are no further attributes in the list.
Attributes may be attached to declarations and to types; these attributes may be accessed with the following macros. All attributes are stored in this way, and many also cause other changes to the declaration or type or to other internal compiler data structures.

\[
\text{tree DECL_ATTRIBUTES (tree decl)}
\]
This macro returns the attributes on the declaration \textit{decl}.

\[
\text{tree TYPE_ATTRIBUTES (tree type)}
\]
This macro returns the attributes on the type \textit{type}.

### 11.6 Expressions

The internal representation for expressions is for the most part quite straightforward. However, there are a few facts that one must bear in mind. In particular, the expression “tree” is actually a directed acyclic graph. (For example there may be many references to the integer constant zero throughout the source program; many of these will be represented by the same expression node.) You should not rely on certain kinds of node being shared, nor should you rely on certain kinds of nodes being unshared.

The following macros can be used with all expression nodes:

\[
\text{TREE_TYPE}
\]
Returns the type of the expression. This value may not be precisely the same type that would be given the expression in the original program.

In what follows, some nodes that one might expect to always have type \texttt{bool} are documented to have either integral or boolean type. At some point in the future, the C front end may also make use of this same intermediate representation, and at this point these nodes will certainly have integral type. The previous sentence is not meant to imply that the C++ front end does not or will not give these nodes integral type.

Below, we list the various kinds of expression nodes. Except where noted otherwise, the operands to an expression are accessed using the \texttt{TREE_OPERAND} macro. For example, to access the first operand to a binary plus expression \texttt{expr}, use:

\[
\text{TREE_OPERAND (expr, 0)}
\]
As this example indicates, the operands are zero-indexed.

#### 11.6.1 Constant expressions

The table below begins with constants, moves on to unary expressions, then proceeds to binary expressions, and concludes with various other kinds of expressions:

\[
\text{INTEGER_CST}
\]
These nodes represent integer constants. Note that the type of these constants is obtained with \texttt{TREE_TYPE}; they are not always of type \texttt{int}. In particular, \texttt{char} constants are represented with \texttt{INTEGER_CST} nodes. The value of the integer constant \texttt{e} is given by

\[
((\text{TREE_INT_CST_HIGH (e)} << \text{HOST_BITS_PER_WIDE_INT}) + \text{TREE_INT_CST_LOW (e)})
\]
\text{HOST_BITS_PER_WIDE_INT} is at least thirty-two on all platforms. Both \texttt{TREE_INT_CST_HIGH} and \texttt{TREE_INT_CST_LOW} return a \texttt{HOST_WIDE_INT}. The
value of an INTEGER_CST is interpreted as a signed or unsigned quantity depending on the type of the constant. In general, the expression given above will overflow, so it should not be used to calculate the value of the constant.

The variable integer_zero_node is an integer constant with value zero. Similarly, integer_one_node is an integer constant with value one. The size_zero_node and size_one_node variables are analogous, but have type size_t rather than int.

The function tree_int_cst_lt is a predicate which holds if its first argument is less than its second. Both constants are assumed to have the same signedness (i.e., either both should be signed or both should be unsigned.) The full width of the constant is used when doing the comparison; the usual rules about promotions and conversions are ignored. Similarly, tree_int_cst_equal holds if the two constants are equal. The tree_int_cst_sgn function returns the sign of a constant. The value is 1, 0, or -1 according on whether the constant is greater than, equal to, or less than zero. Again, the signedness of the constant’s type is taken into account; an unsigned constant is never less than zero, no matter what its bit-pattern.

REAL_CST

FIXME: Talk about how to obtain representations of this constant, do comparisons, and so forth.

FIXED_CST

These nodes represent fixed-point constants. The type of these constants is obtained with TREE_TYPE. TREE_FIXED_CST_PTR points to a struct fixed_value; TREE_FIXED_CST returns the structure itself. struct fixed_value contains data with the size of two HOST_BITS_PER_WIDE_INT and mode as the associated fixed-point machine mode for data.

COMPLEX_CST

These nodes are used to represent complex number constants, that is a __complex__ whose parts are constant nodes. The TREE_REALPART and TREE_IMAGPART return the real and the imaginary parts respectively.

VECTOR_CST

These nodes are used to represent vector constants, whose parts are constant nodes. Each individual constant node is either an integer or a double constant node. The first operand is a TREE_LIST of the constant nodes and is accessed through TREE_VECTOR_CST_ELTS.

STRING_CST

These nodes represent string-constants. The TREE_STRING_LENGTH returns the length of the string, as an int. The TREE_STRING_POINTER is a char* containing the string itself. The string may not be NUL-terminated, and it may contain embedded NUL characters. Therefore, the TREE_STRING_LENGTH includes the trailing NUL if it is present.

For wide string constants, the TREE_STRING_LENGTH is the number of bytes in the string, and the TREE_STRING_POINTER points to an array of the bytes of the string, as represented on the target system (that is, as integers in the target
endianness). Wide and non-wide string constants are distinguished only by the TREE_TYPE of the STRING_CST.

FIXME: The formats of string constants are not well-defined when the target system bytes are not the same width as host system bytes.

### 11.6.2 References to storage

**ARRAY_REF**

These nodes represent array accesses. The first operand is the array; the second is the index. To calculate the address of the memory accessed, you must scale the index by the size of the type of the array elements. The type of these expressions must be the type of a component of the array. The third and fourth operands are used after gimplification to represent the lower bound and component size but should not be used directly; call array_ref_low_bound and array_ref_element_size instead.

**ARRAY_RANGE_REF**

These nodes represent access to a range (or “slice”) of an array. The operands are the same as that for ARRAY_REF and have the same meanings. The type of these expressions must be an array whose component type is the same as that of the first operand. The range of that array type determines the amount of data these expressions access.

**TARGET_MEM_REF**

These nodes represent memory accesses whose address directly map to an addressing mode of the target architecture. The first argument is TMR_SYMBOL and must be a VAR_DECL of an object with a fixed address. The second argument is TMR_BASE and the third one is TMR_INDEX. The fourth argument is TMR_STEP and must be an INTEGER_CST. The fifth argument is TMR_OFFSET and must be an INTEGER_CST. Any of the arguments may be NULL if the appropriate component does not appear in the address. Address of the TARGET_MEM_REF is determined in the following way.

\[ \&\text{TMR_SYMBOL} + \text{TMR_BASE} + \text{TMR_INDEX} \times \text{TMR_STEP} + \text{TMR_OFFSET} \]

The sixth argument is the reference to the original memory access, which is preserved for the purposes of the RTL alias analysis. The seventh argument is a tag representing the results of tree level alias analysis.

**ADDR_EXPR**

These nodes are used to represent the address of an object. (These expressions will always have pointer or reference type.) The operand may be another expression, or it may be a declaration.

As an extension, GCC allows users to take the address of a label. In this case, the operand of the ADDR_EXPR will be a LABELDECL. The type of such an expression is void*.

If the object addressed is not an lvalue, a temporary is created, and the address of the temporary is used.
**INDIRECT_REF**
These nodes are used to represent the object pointed to by a pointer. The operand is the pointer being dereferenced; it will always have pointer or reference type.

**MEM_REF**
These nodes are used to represent the object pointed to by a pointer offset by a constant. The first operand is the pointer being dereferenced; it will always have pointer or reference type. The second operand is a pointer constant. Its type is specifying the type to be used for type-based alias analysis.

**COMPONENT_REF**
These nodes represent non-static data member accesses. The first operand is the object (rather than a pointer to it); the second operand is the FIELD_DECL for the data member. The third operand represents the byte offset of the field, but should not be used directly; call component_ref_field_offset instead.

### 11.6.3 Unary and Binary Expressions

**NEGATE_EXPR**
These nodes represent unary negation of the single operand, for both integer and floating-point types. The type of negation can be determined by looking at the type of the expression.

The behavior of this operation on signed arithmetic overflow is controlled by the **flag_wrapv** and **flag_trapv** variables.

**ABS_EXPR**
These nodes represent the absolute value of the single operand, for both integer and floating-point types. This is typically used to implement the abs, labs and llabs builtins for integer types, and the fabs, fabsf and fabsl builtins for floating point types. The type of abs operation can be determined by looking at the type of the expression.

This node is not used for complex types. To represent the modulus or complex abs of a complex value, use the **BUILT_IN_CABS**, **BUILT_IN_CABSF** or **BUILT_IN_CABSL** builtins, as used to implement the C99 cabs, cabsf and cabsl built-in functions.

**BIT_NOT_EXPR**
These nodes represent bitwise complement, and will always have integral type. The only operand is the value to be complemented.

**TRUTH_NOT_EXPR**
These nodes represent logical negation, and will always have integral (or boolean) type. The operand is the value being negated. The type of the operand and that of the result are always of **BOOLEAN_TYPE** or **INTEGER_TYPE**.

**PREDECREMENT_EXPR**
**PREINCREMENT_EXPR**
**POSTDECREMENT_EXPR**
**POSTINCREMENT_EXPR**
These nodes represent increment and decrement expressions. The value of the single operand is computed, and the operand incremented or decremented. In
the case of PREDECREMENT_EXPR and PREINCREMENT_EXPR, the value of the expression is the value resulting after the increment or decrement; in the case of POSTDECREMENT_EXPR and POSTINCREMENT_EXPR is the value before the increment or decrement occurs. The type of the operand, like that of the result, will be either integral, boolean, or floating-point.

FIX_TRUNC_EXPR

These nodes represent conversion of a floating-point value to an integer. The single operand will have a floating-point type, while the complete expression will have an integral (or boolean) type. The operand is rounded towards zero.

FLOAT_EXPR

These nodes represent conversion of an integral (or boolean) value to a floating-point value. The single operand will have integral type, while the complete expression will have a floating-point type.

FIXME: How is the operand supposed to be rounded? Is this dependent on ‘-mieee’?

COMPLEX_EXPR

These nodes are used to represent complex numbers constructed from two expressions of the same (integer or real) type. The first operand is the real part and the second operand is the imaginary part.

CONJ_EXPR

These nodes represent the conjugate of their operand.

REALPART_EXPR

IMAGPART_EXPR

These nodes represent respectively the real and the imaginary parts of complex numbers (their sole argument).

NON_LVALUE_EXPR

These nodes indicate that their one and only operand is not an lvalue. A back end can treat these identically to the single operand.

NOP_EXPR

These nodes are used to represent conversions that do not require any code-generation. For example, conversion of a char* to an int* does not require any code be generated; such a conversion is represented by a NOP_EXPR. The single operand is the expression to be converted. The conversion from a pointer to a reference is also represented with a NOP_EXPR.

CONVERT_EXPR

These nodes are similar to NOP_EXPRs, but are used in those situations where code may need to be generated. For example, if an int* is converted to an int code may need to be generated on some platforms. These nodes are never used for C++-specific conversions, like conversions between pointers to different classes in an inheritance hierarchy. Any adjustments that need to be made in such cases are always indicated explicitly. Similarly, a user-defined conversion is never represented by a CONVERT_EXPR; instead, the function calls are made explicit.
These nodes are used to represent conversions that involve fixed-point values. For example, from a fixed-point value to another fixed-point value, from an integer to a fixed-point value, from a fixed-point value to an integer, from a floating-point value to a fixed-point value, or from a fixed-point value to a floating-point value.

These nodes represent left and right shifts, respectively. The first operand is the value to shift; it will always be of integral type. The second operand is an expression for the number of bits by which to shift. Right shift should be treated as arithmetic, i.e., the high-order bits should be zero-filled when the expression has unsigned type and filled with the sign bit when the expression has signed type. Note that the result is undefined if the second operand is larger than or equal to the first operand’s type size.

These nodes represent bitwise inclusive or, bitwise exclusive or, and bitwise and, respectively. Both operands will always have integral type.

These nodes represent logical “and” and logical “or”, respectively. These operators are not strict; i.e., the second operand is evaluated only if the value of the expression is not determined by evaluation of the first operand. The type of the operands and that of the result are always of BOOLEAN_TYPE or INTEGER_TYPE.

These nodes represent logical and, logical or, and logical exclusive or. They are strict; both arguments are always evaluated. There are no corresponding operators in C or C++, but the front end will sometimes generate these expressions anyhow, if it can tell that strictness does not matter. The type of the operands and that of the result are always of BOOLEAN_TYPE or INTEGER_TYPE.

This node represents pointer arithmetic. The first operand is always a pointer/reference type. The second operand is always an unsigned integer type compatible with sizetype. This is the only binary arithmetic operand that can operate on pointer types.

These nodes represent various binary arithmetic operations. Respectively, these operations are addition, subtraction (of the second operand from the first) and multiplication. Their operands may have either integral or floating type, but
there will never be case in which one operand is of floating type and the other is of integral type.
The behavior of these operations on signed arithmetic overflow is controlled by the \texttt{flag\_wrapv} and \texttt{flag\_trapv} variables.

\textbf{RDIV\_EXPR}
This node represents a floating point division operation.

\textbf{TRUNC\_DIV\_EXPR}
\textbf{FLOOR\_DIV\_EXPR}
\textbf{CEIL\_DIV\_EXPR}
\textbf{ROUND\_DIV\_EXPR}
These nodes represent integer division operations that return an integer result. \texttt{TRUNC\_DIV\_EXPR} rounds towards zero, \texttt{FLOOR\_DIV\_EXPR} rounds towards negative infinity, \texttt{CEIL\_DIV\_EXPR} rounds towards positive infinity and \texttt{ROUND\_DIV\_EXPR} rounds to the closest integer. Integer division in C and C++ is truncating, i.e. \texttt{TRUNC\_DIV\_EXPR}.

The behavior of these operations on signed arithmetic overflow, when dividing the minimum signed integer by minus one, is controlled by the \texttt{flag\_wrapv} and \texttt{flag\_trapv} variables.

\textbf{TRUNC\_MOD\_EXPR}
\textbf{FLOOR\_MOD\_EXPR}
\textbf{CEIL\_MOD\_EXPR}
\textbf{ROUND\_MOD\_EXPR}
These nodes represent the integer remainder or modulus operation. The integer modulus of two operands \texttt{a} and \texttt{b} is defined as \(a - (a/b)\times b\) where the division calculated using the corresponding division operator. Hence for \texttt{TRUNC\_MOD\_EXPR} this definition assumes division using truncation towards zero, i.e. \texttt{TRUNC\_DIV\_EXPR}. Integer remainder in C and C++ uses truncating division, i.e. \texttt{TRUNC\_MOD\_EXPR}.

\textbf{EXACT\_DIV\_EXPR}
The \texttt{EXACT\_DIV\_EXPR} code is used to represent integer divisions where the numerator is known to be an exact multiple of the denominator. This allows the backend to choose between the faster of \texttt{TRUNC\_DIV\_EXPR}, \texttt{CEIL\_DIV\_EXPR} and \texttt{FLOOR\_DIV\_EXPR} for the current target.

\textbf{LT\_EXPR}
\textbf{LE\_EXPR}
\textbf{GT\_EXPR}
\textbf{GE\_EXPR}
\textbf{EQ\_EXPR}
\textbf{NE\_EXPR}
These nodes represent the less than, less than or equal to, greater than, greater than or equal to, equal, and not equal comparison operators. The first and second operand with either be both of integral type or both of floating type. The result type of these expressions will always be of integral or boolean type. These operations return the result type’s zero value for false, and the result type’s one value for true.
For floating point comparisons, if we honor IEEE NaNs and either operand is NaN, then `NE_EXPR` always returns true and the remaining operators always return false. On some targets, comparisons against an IEEE NaN, other than equality and inequality, may generate a floating point exception.

**ORDERED_EXPR**
**UNORDERED_EXPR**

These nodes represent non-trapping ordered and unordered comparison operators. These operations take two floating point operands and determine whether they are ordered or unordered relative to each other. If either operand is an IEEE NaN, their comparison is defined to be unordered, otherwise the comparison is defined to be ordered. The result type of these expressions will always be of integral or boolean type. These operations return the result type’s zero value for false, and the result type’s one value for true.

**UNLT_EXPR**
**UNLE_EXPR**
**UNGT_EXPR**
**UNGE_EXPR**
**UNEQ_EXPR**
**LTGT_EXPR**

These nodes represent the unordered comparison operators. These operations take two floating point operands and determine whether the operands are unordered or are less than, less than or equal to, greater than, greater than or equal to, or equal respectively. For example, `UNLT_EXPR` returns true if either operand is an IEEE NaN or the first operand is less than the second. With the possible exception of `LTGT_EXPR`, all of these operations are guaranteed not to generate a floating point exception. The result type of these expressions will always be of integral or boolean type. These operations return the result type’s zero value for false, and the result type’s one value for true.

**MODIFY_EXPR**

These nodes represent assignment. The left-hand side is the first operand; the right-hand side is the second operand. The left-hand side will be a `VAR_DECL`, `INDIRECT_REF`, `COMPONENT_REF`, or other lvalue.

These nodes are used to represent not only assignment with ‘=’ but also compound assignments (like ‘+=’), by reduction to ‘=’ assignment. In other words, the representation for ‘i += 3’ looks just like that for ‘i = i + 3’.

**INIT_EXPR**

These nodes are just like `MODIFY_EXPR`, but are used only when a variable is initialized, rather than assigned to subsequently. This means that we can assume that the target of the initialization is not used in computing its own value; any reference to the lhs in computing the rhs is undefined.

**COMPOUND_EXPR**

These nodes represent comma-expressions. The first operand is an expression whose value is computed and thrown away prior to the evaluation of the second operand. The value of the entire expression is the value of the second operand.
COND_EXPR

These nodes represent ?: expressions. The first operand is of boolean or integral type. If it evaluates to a nonzero value, the second operand should be evaluated, and returned as the value of the expression. Otherwise, the third operand is evaluated, and returned as the value of the expression.

The second operand must have the same type as the entire expression, unless it unconditionally throws an exception or calls a noreturn function, in which case it should have void type. The same constraints apply to the third operand. This allows array bounds checks to be represented conveniently as (i >= 0 && i < 10) ? i : abort().

As a GNU extension, the C language front-ends allow the second operand of the ?: operator may be omitted in the source. For example, x ? : 3 is equivalent to x ? x : 3, assuming that x is an expression without side-effects. In the tree representation, however, the second operand is always present, possibly protected by SAVE_EXPR if the first argument does cause side-effects.

CALL_EXPR

These nodes are used to represent calls to functions, including non-static member functions. CALL_EXPRs are implemented as expression nodes with a variable number of operands. Rather than using TREE_OPERAND to extract them, it is preferable to use the specialized accessor macros and functions that operate specifically on CALL_EXPR nodes.

CALL_EXPR_FN returns a pointer to the function to call; it is always an expression whose type is a POINTER_TYPE.

The number of arguments to the call is returned by call_expr_nargs, while the arguments themselves can be accessed with the CALL_EXPR_ARG macro. The arguments are zero-indexed and numbered left-to-right. You can iterate over the arguments using FOR_EACH_CALL_EXPR_ARG, as in:

```c
    tree call, arg;
    call_expr_arg_iterator iter;
    FOR_EACH_CALL_EXPR_ARG (arg, iter, call)
        /* arg is bound to successive arguments of call. */
    ...;
```

For non-static member functions, there will be an operand corresponding to the this pointer. There will always be expressions corresponding to all of the arguments, even if the function is declared with default arguments and some arguments are not explicitly provided at the call sites.

CALL_EXPRs also have a CALL_EXPR_STATIC_CHAIN operand that is used to implement nested functions. This operand is otherwise null.

CLEANUP_POINT_EXPR

These nodes represent full-expressions. The single operand is an expression to evaluate. Any destructor calls engendered by the creation of temporaries during the evaluation of that expression should be performed immediately after the expression is evaluated.

CONSTRUCTOR

These nodes represent the brace-enclosed initializers for a structure or array. The first operand is reserved for use by the back end. The second operand
is a TREE_LIST. If the TREE_TYPE of the CONSTRUCTOR is a RECORD_TYPE or
UNION_TYPE, then the TREE_PURPOSE of each node in the TREE_LIST will be a
FIELDDECL and the TREE_VALUE of each node will be the expression used to
initialize that field.

If the TREE_TYPE of the CONSTRUCTOR is an ARRAY_TYPE, then the TREE_PURPOSE
of each element in the TREE_LIST will be an INTEGER_CST or a RANGE_EXPR
of two INTEGER_CSTs. A single INTEGER_CST indicates which element of the
array (indexed from zero) is being assigned to. A RANGE_EXPR indicates an
inclusive range of elements to initialize. In both cases the TREE_VALUE is the
corresponding initializer. It is re-evaluated for each element of a RANGE_EXPR.
If the TREE_PURPOSE is NULL_TREE, then the initializer is for the next available
array element.

In the front end, you should not depend on the fields appearing in any particular
order. However, in the middle end, fields must appear in declaration order. You
should not assume that all fields will be represented. Unrepresented fields will
be set to zero.

COMPOUND_LITERAL_EXPR
These nodes represent ISO C99 compound literals. The COMPOUND_LITERAL_
EXPR_DECL_EXPR is a DECL_EXPR containing an anonymous VAR_DECL for the
unnamed object represented by the compound literal; the DECL_INITIAL of that
VAR_DECL is a CONSTRUCTOR representing the brace-enclosed list of initializers in
the compound literal. That anonymous VAR_DECL can also be accessed directly
by the COMPOUND_LITERAL_EXPR_DECL macro.

SAVE_EXPR
A SAVE_EXPR represents an expression (possibly involving side-effects) that is
used more than once. The side-effects should occur only the first time the
expression is evaluated. Subsequent uses should just reuse the computed value.
The first operand to the SAVE_EXPR is the expression to evaluate. The side-
effects should be executed where the SAVE_EXPR is first encountered in a depth-
first preorder traversal of the expression tree.

TARGET_EXPR
A TARGET_EXPR represents a temporary object. The first operand is a VAR_ DECL for the temporary variable. The second operand is the initializer for the
temporary. The initializer is evaluated and, if non-void, copied (bitwise) into
the temporary. If the initializer is void, that means that it will perform the
initialization itself.

Often, a TARGET_EXPR occurs on the right-hand side of an assignment, or as
the second operand to a comma-expression which is itself the right-hand side
of an assignment, etc. In this case, we say that the TARGET_EXPR is “normal”;
otherwise, we say it is “orphaned”. For a normal TARGET_EXPR the temporary
variable should be treated as an alias for the left-hand side of the assignment,
rather than as a new temporary variable.

The third operand to the TARGET_EXPR, if present, is a cleanup-expression (i.e.,
destructor call) for the temporary. If this expression is orphaned, then this
expression must be executed when the statement containing this expression is
complete. These cleanups must always be executed in the order opposite to that in which they were encountered. Note that if a temporary is created on one branch of a conditional operator (i.e., in the second or third operand to a `COND_EXPR`), the cleanup must be run only if that branch is actually executed.

**VA_ARG_EXPR**

This node is used to implement support for the C/C++ variable argument-list mechanism. It represents expressions like `va_arg (ap, type)`. Its `TREE_TYPE` yields the tree representation for `type` and its sole argument yields the representation for `ap`.

### 11.6.4 Vectors

**VEC_LSHIFT_EXPR**

**VEC_RSHIFT_EXPR**

These nodes represent whole vector left and right shifts, respectively. The first operand is the vector to shift; it will always be of vector type. The second operand is an expression for the number of bits by which to shift. Note that the result is undefined if the second operand is larger than or equal to the first operand’s type size.

**VEC_WIDEN_MULT_HI_EXPR**

**VEC_WIDEN_MULT_LO_EXPR**

These nodes represent widening vector multiplication of the high and low parts of the two input vectors, respectively. Their operands are vectors that contain the same number of elements (N) of the same integral type. The result is a vector that contains half as many elements, of an integral type whose size is twice as wide. In the case of `VEC_WIDEN_MULT_HI_EXPR` the high N/2 elements of the two vector are multiplied to produce the vector of N/2 products. In the case of `VEC_WIDEN_MULT_LO_EXPR` the low N/2 elements of the two vector are multiplied to produce the vector of N/2 products.

**VEC_UNPACK_HI_EXPR**

**VEC_UNPACK_LO_EXPR**

These nodes represent unpacking of the high and low parts of the input vector, respectively. The single operand is a vector that contains N elements of the same integral or floating point type. The result is a vector that contains half as many elements, of an integral or floating point type whose size is twice as wide. In the case of `VEC_UNPACK_HI_EXPR` the high N/2 elements of the vector are extracted and widened (promoted). In the case of `VEC_UNPACK_LO_EXPR` the low N/2 elements of the vector are extracted and widened (promoted).

**VEC_UNPACK_FLOAT_HI_EXPR**

**VEC_UNPACK_FLOAT_LO_EXPR**

These nodes represent unpacking of the high and low parts of the input vector, where the values are converted from fixed point to floating point. The single operand is a vector that contains N elements of the same integral type. The result is a vector that contains half as many elements of a floating point type whose size is twice as wide. In the case of `VEC_UNPACK_HI_EXPR` the high N/2
elements of the vector are extracted, converted and widened. In the case of \texttt{VEC\_UNPACK\_LO\_EXPR} the low \(N/2\) elements of the vector are extracted, converted and widened.

\texttt{VEC\_PACK\_TRUNC\_EXPR}

This node represents packing of truncated elements of the two input vectors into the output vector. Input operands are vectors that contain the same number of elements of the same integral or floating point type. The result is a vector that contains twice as many elements of an integral or floating point type whose size is half as wide. The elements of the two vectors are demoted and merged (concatenated) to form the output vector.

\texttt{VEC\_PACK\_SAT\_EXPR}

This node represents packing of elements of the two input vectors into the output vector using saturation. Input operands are vectors that contain the same number of elements of the same integral type. The result is a vector that contains twice as many elements of an integral type whose size is half as wide. The elements of the two vectors are demoted and merged (concatenated) to form the output vector.

\texttt{VEC\_PACK\_FIX\_TRUNC\_EXPR}

This node represents packing of elements of the two input vectors into the output vector, where the values are converted from floating point to fixed point. Input operands are vectors that contain the same number of elements of a floating point type. The result is a vector that contains twice as many elements of an integral type whose size is half as wide. The elements of the two vectors are merged (concatenated) to form the output vector.

\texttt{VEC\_EXTRACT\_EVEN\_EXPR}
\texttt{VEC\_EXTRACT\_ODD\_EXPR}

These nodes represent extracting of the even/odd elements of the two input vectors, respectively. Their operands and result are vectors that contain the same number of elements of the same type.

\texttt{VEC\_INTERLEAVE\_HIGH\_EXPR}
\texttt{VEC\_INTERLEAVE\_LOW\_EXPR}

These nodes represent merging and interleaving of the high/low elements of the two input vectors, respectively. The operands and the result are vectors that contain the same number of elements \((N)\) of the same type. In the case of \texttt{VEC\_INTERLEAVE\_HIGH\_EXPR}, the high \(N/2\) elements of the first input vector are interleaved with the high \(N/2\) elements of the second input vector. In the case of \texttt{VEC\_INTERLEAVE\_LOW\_EXPR}, the low \(N/2\) elements of the first input vector are interleaved with the low \(N/2\) elements of the second input vector.

11.7 Statements

Most statements in GIMPLE are assignment statements, represented by \texttt{GIMPLE\_ASSIGN}. No other C expressions can appear at statement level; a reference to a volatile object is converted into a \texttt{GIMPLE\_ASSIGN}.

There are also several varieties of complex statements.
11.7.1 Basic Statements

**ASM_EXPR**

Used to represent an inline assembly statement. For an inline assembly statement like:

```c
asm ("mov x, y");
```

The `ASM_STRING` macro will return a `STRING_CST` node for "mov x, y". If the original statement made use of the extended-assembly syntax, then `ASM_OUTPUTS`, `ASM_INPUTS`, and `ASM_CLOBBERS` will be the outputs, inputs, and clobbers for the statement, represented as `STRING_CST` nodes. The extended-assembly syntax looks like:

```c
asm ("fsinx %1,%0" : "=f" (result) : "f" (angle));
```

The first string is the `ASM_STRING`, containing the instruction template. The next two strings are the output and inputs, respectively; this statement has no clobbers. As this example indicates, “plain” assembly statements are merely a special case of extended assembly statements; they have no cv-qualifiers, outputs, inputs, or clobbers. All of the strings will be NUL-terminated, and will contain no embedded NUL-characters.

If the assembly statement is declared `volatile`, or if the statement was not an extended assembly statement, and is therefore implicitly volatile, then the predicate `ASM_VOLATILE_P` will hold of the `ASM_EXPR`.

**DECL_EXPR**

Used to represent a local declaration. The `DECL_EXPR_DECL` macro can be used to obtain the entity declared. This declaration may be a `LABEL_DECL`, indicating that the label declared is a local label. (As an extension, GCC allows the declaration of labels with scope.) In C, this declaration may be a `FUNCTION_DECL`, indicating the use of the GCC nested function extension. For more information, see Section 11.8 [Functions], page 206.

**LABEL_EXPR**

Used to represent a label. The `LABEL_DECL` declared by this statement can be obtained with the `LABEL_EXPR_LABEL` macro. The `IDENTIFIER_NODE` giving the name of the label can be obtained from the `LABEL_DECL` with `DECL_NAME`.

**GOTO_EXPR**

Used to represent a `goto` statement. The `GOTO_DESTINATION` will usually be a `LABEL_DECL`. However, if the “computed goto” extension has been used, the `GOTO_DESTINATION` will be an arbitrary expression indicating the destination. This expression will always have pointer type.

**RETURN_EXPR**

Used to represent a `return` statement. Operand 0 represents the value to return. It should either be the `RESULT_DECL` for the containing function, or a `MODIFY_EXPR` or `INIT_EXPR` setting the function’s `RESULT_DECL`. It will be `NULL_TREE` if the statement was just

```c
    return;
```
LOOP_EXPR
These nodes represent “infinite” loops. The LOOP_EXPR_BODY represents the body of the loop. It should be executed forever, unless an EXIT_EXPR is encountered.

EXIT_EXPR
These nodes represent conditional exits from the nearest enclosing LOOP_EXPR. The single operand is the condition; if it is nonzero, then the loop should be exited. An EXIT_EXPR will only appear within a LOOP_EXPR.

SWITCH_STMT
Used to represent a switch statement. The SWITCH_STMT_COND is the expression on which the switch is occurring. See the documentation for an IF_STMT for more information on the representation used for the condition. The SWITCH_STMT_BODY is the body of the switch statement. The SWITCH_STMT_TYPE is the original type of switch expression as given in the source, before any compiler conversions.

CASE_LABEL_EXPR
Use to represent a case label, range of case labels, or a default label. If CASE_LOW is NULL_TREE, then this is a default label. Otherwise, if CASE_HIGH is NULL_TREE, then this is an ordinary case label. In this case, CASE_LOW is an expression giving the value of the label. Both CASE_LOW and CASE_HIGH are INTEGER_CST nodes. These values will have the same type as the condition expression in the switch statement.

Otherwise, if both CASE_LOW and CASE_HIGH are defined, the statement is a range of case labels. Such statements originate with the extension that allows users to write things of the form:

    case 2 ... 5:

The first value will be CASE_LOW, while the second will be CASE_HIGH.

11.7.2 Blocks
Block scopes and the variables they declare in GENERIC are expressed using the BIND_EXPR code, which in previous versions of GCC was primarily used for the C statement-expression extension.

Variables in a block are collected into BIND_EXPR_VARS in declaration order through their TREE_CHAIN field. Any runtime initialization is moved out of DECL_INITIAL and into a statement in the controlled block. When simplifying from C or C++, this initialization replaces the DECL_STMT. These variables will never require cleanups. The scope of these variables is just the body.

Variable-length arrays (VLAs) complicate this process, as their size often refers to variables initialized earlier in the block. To handle this, we currently split the block at that point, and move the VLA into a new, inner BIND_EXPR. This strategy may change in the future.

A C++ program will usually contain more BIND_EXPRs than there are syntactic blocks in the source code, since several C++ constructs have implicit scopes associated with them. On the other hand, although the C++ front end uses pseudo-scopes to handle cleanups for
objects with destructors, these don’t translate into the GIMPLE form; multiple declarations
at the same level use the same BIND_EXPR.

11.7.3 Statement Sequences
Multiple statements at the same nesting level are collected into a STATEMENT_LIST. Statement
lists are modified and traversed using the interface in ‘tree-iterator.h’.

11.7.4 Empty Statements
Whenever possible, statements with no effect are discarded. But if they are nested within
another construct which cannot be discarded for some reason, they are instead replaced
with an empty statement, generated by build_empty_stmt. Initially, all empty statements
were shared, after the pattern of the Java front end, but this caused a lot of trouble in
practice.

An empty statement is represented as (void)0.

11.7.5 Jumps
Other jumps are expressed by either GOTO_EXPR or RETURN_EXPR.

The operand of a GOTO_EXPR must be either a label or a variable containing the address
to jump to.

The operand of a RETURN_EXPR is either NULL_TREE, RESULT_DECL, or a MODIFY_EXPR
which sets the return value. It would be nice to move the MODIFY_EXPR into a separate
statement, but the special return semantics in expand_return make that difficult. It may
still happen in the future, perhaps by moving most of that logic into expand_assignment.

11.7.6 Cleanups
Destructors for local C++ objects and similar dynamic cleanups are represented in GIM-
PLE by a TRY_FINALLY_EXPR. TRY_FINALLY_EXPR has two operands, both of which are a
sequence of statements to execute. The first sequence is executed. When it completes the
second sequence is executed.

The first sequence may complete in the following ways:
1. Execute the last statement in the sequence and fall off the end.
2. Execute a goto statement (GOTO_EXPR) to an ordinary label outside the sequence.
3. Execute a return statement (RETURN_EXPR).
4. Throw an exception. This is currently not explicitly represented in GIMPLE.

The second sequence is not executed if the first sequence completes by calling setjmp or
exit or any other function that does not return. The second sequence is also not executed
if the first sequence completes via a non-local goto or a computed goto (in general the
compiler does not know whether such a goto statement exits the first sequence or not, so
we assume that it doesn’t).

After the second sequence is executed, if it completes normally by falling off the end,
execution continues wherever the first sequence would have continued, by falling off the
end, or doing a goto, etc.

TRY_FINALLY_EXPR complicates the flow graph, since the cleanup needs to appear on
every edge out of the controlled block; this reduces the freedom to move code across these
edges. Therefore, the EH lowering pass which runs before most of the optimization passes eliminates these expressions by explicitly adding the cleanup to each edge. Rethrowing the exception is represented using `RESX_EXPR`.

### 11.7.7 OpenMP

All the statements starting with `OMP_` represent directives and clauses used by the OpenMP API [http://www.openmp.org/](http://www.openmp.org/).

**OMP_PARALLEL**

Represents `#pragma omp parallel [clause1 ... clauseN]`. It has four operands:
- **Operand OMP_PARALLEL_BODY** is valid while in GENERIC and High GIMPLE forms. It contains the body of code to be executed by all the threads. During GIMPLE lowering, this operand becomes `NULL` and the body is emitted linearly after `OMP_PARALLEL`.
- **Operand OMP_PARALLEL_CLAUSES** is the list of clauses associated with the directive.
- **Operand OMP_PARALLEL_FN** is created by `pass_lower_omp`, it contains the `FUNCTION_DECL` for the function that will contain the body of the parallel region.
- **Operand OMP_PARALLEL_DATA_ARG** is also created by `pass_lower_omp`. If there are shared variables to be communicated to the children threads, this operand will contain the `VAR_DECL` that contains all the shared values and variables.

**OMP_FOR**

Represents `#pragma omp for [clause1 ... clauseN]`. It has 5 operands:
- **Operand OMP_FOR_BODY** contains the loop body.
- **Operand OMP_FOR_CLAUSES** is the list of clauses associated with the directive.
- **Operand OMP_FOR_INIT** is the loop initialization code of the form `VAR = N1`.
- **Operand OMP_FOR_COND** is the loop conditional expression of the form `VAR {<,>,<,>,>=} N2`.
- **Operand OMP_FOR_INCR** is the loop index increment of the form `VAR {+=,-=} INCR`.
- **Operand OMP_FOR_PRE_BODY** contains side-effect code from operands `OMP_FOR_INIT`, `OMP_FOR_COND` and `OMP_FOR_INCR`. These side-effects are part of the `OMP_FOR` block but must be evaluated before the start of loop body.
- The loop index variable `VAR` must be a signed integer variable, which is implicitly private to each thread. Bounds `N1` and `N2` and the increment expression `INCR` are required to be loop invariant integer expressions that are evaluated without any synchronization. The evaluation order, frequency of evaluation and side-effects are unspecified by the standard.

**OMP_SECTIONS**

Represents `#pragma omp sections [clause1 ... clauseN]`.
- **Operand OMP_SECTIONS_BODY** contains the sections body, which in turn contains a set of `OMP_SECTION` nodes for each of the concurrent sections delimited by `#pragma omp section`. 

Operand **OMP_SECTIONS_CLAUSES** is the list of clauses associated with the directive.

**OMP_SECTION**
Section delimiter for **OMP_SECTIONS**.

**OMP_SINGLE**
Represents **#pragma omp single**.
Operand **OMP_SINGLE_BODY** contains the body of code to be executed by a single thread.
Operand **OMP_SINGLE_CLAUSES** is the list of clauses associated with the directive.

**OMP_MASTER**
Represents **#pragma omp master**.
Operand **OMP_MASTER_BODY** contains the body of code to be executed by the master thread.

**OMP_ORDERED**
Represents **#pragma omp ordered**.
Operand **OMP_ORDERED_BODY** contains the body of code to be executed in the sequential order dictated by the loop index variable.

**OMP_CRITICAL**
Represents **#pragma omp critical [name]**.
Operand **OMP_CRITICAL_BODY** is the critical section.
Operand **OMP_CRITICAL_NAME** is an optional identifier to label the critical section.

**OMP_RETURN**
This does not represent any OpenMP directive, it is an artificial marker to indicate the end of the body of an OpenMP. It is used by the flow graph (**tree-cfg.c**) and OpenMP region building code (**omp-low.c**).

**OMP_CONTINUE**
Similarly, this instruction does not represent an OpenMP directive, it is used by **OMP_FOR** and **OMP_SECTIONS** to mark the place where the code needs to loop to the next iteration (in the case of **OMP_FOR**) or the next section (in the case of **OMP_SECTIONS**).
In some cases, **OMP_CONTINUE** is placed right before **OMP_RETURN**. But if there are cleanups that need to occur right after the looping body, it will be emitted between **OMP_CONTINUE** and **OMP_RETURN**.

**OMP_ATOMIC**
Represents **#pragma omp atomic**.
Operand 0 is the address at which the atomic operation is to be performed.
Operand 1 is the expression to evaluate. The gimplifier tries three alternative code generation strategies. Whenever possible, an atomic update built-in is used. If that fails, a compare-and-swap loop is attempted. If that also fails, a regular critical section around the expression is used.
OMP_CLAUSE

Represents clauses associated with one of the OMP_ directives. Clauses are represented by separate sub-codes defined in ‘tree.h’. Clauses codes can be one of: OMP_CLAUSE_PRIVATE, OMP_CLAUSE_SHARED, OMP_CLAUSE_FIRSTPRIVATE, OMP_CLAUSE_LASTPRIVATE, OMP_CLAUSE_COPYIN, OMP_CLAUSE_COPYPRIVATE, OMP_CLAUSE_IF, OMP_CLAUSE_NUM_THREADS, OMP_CLAUSE_SCHEDULE, OMP_CLAUSE_NOWAIT, OMP_CLAUSE_ORDERED, OMP_CLAUSE_DEFAULT, and OMP_CLAUSE_REDUCTION. Each code represents the corresponding OpenMP clause.

Clauses associated with the same directive are chained together via OMP_CLAUSE_CHAIN. Those clauses that accept a list of variables are restricted to exactly one, accessed with OMP_CLAUSE_VAR. Therefore, multiple variables under the same clause C need to be represented as multiple C clauses chained together. This facilitates adding new clauses during compilation.

11.8 Functions

A function is represented by a FUNCTION_DECL node. It stores the basic pieces of the function such as body, parameters, and return type as well as information on the surrounding context, visibility, and linkage.

11.8.1 Function Basics

A function has four core parts: the name, the parameters, the result, and the body. The following macros and functions access these parts of a FUNCTION_DECL as well as other basic features:

DECL_NAME

This macro returns the unqualified name of the function, as an IDENTIFIER_NODE. For an instantiation of a function template, the DECL_NAME is the unqualified name of the template, not something like f<int>. The value of DECL_NAME is undefined when used on a constructor, destructor, overloaded operator, or type-conversion operator, or any function that is implicitly generated by the compiler. See below for macros that can be used to distinguish these cases.

DECL_ASSEMBLER_NAME

This macro returns the mangled name of the function, also an IDENTIFIER_NODE. This name does not contain leading underscores on systems that prefix all identifiers with underscores. The mangled name is computed in the same way on all platforms; if special processing is required to deal with the object file format used on a particular platform, it is the responsibility of the back end to perform those modifications. (Of course, the back end should not modify DECL_ASSEMBLER_NAME itself.)

Using DECL_ASSEMBLER_NAME will cause additional memory to be allocated (for the mangled name of the entity) so it should be used only when emitting assembly code. It should not be used within the optimizers to determine whether or not two declarations are the same, even though some of the existing optimizers do use it in that way. These uses will be removed over time.
DECL_ARGUMENTS
This macro returns the PARM_DECL for the first argument to the function. Subsequent PARM_DECL nodes can be obtained by following the TREE_CHAIN links.

DECL_RESULT
This macro returns the RESULT_DECL for the function.

DECL_SAVED_TREE
This macro returns the complete body of the function.

TREE_TYPE
This macro returns the FUNCTION_TYPE or METHOD_TYPE for the function.

DECL_INITIAL
A function that has a definition in the current translation unit will have a non-NULL DECL_INITIAL. However, back ends should not make use of the particular value given by DECL_INITIAL.

It should contain a tree of BLOCK nodes that mirrors the scopes that variables are bound in the function. Each block contains a list of decls declared in a basic block, a pointer to a chain of blocks at the next lower scope level, then a pointer to the next block at the same level and a backpointer to the parent BLOCK or FUNCTION_DECL. So given a function as follows:

```c
void foo()
{
    int a;
    {
        int b;
    }
    int c;
}
```

you would get the following:

```c
tree foo = FUNCTION_DECL;
tree decl_a = VAR_DECL;
tree decl_b = VAR_DECL;
tree decl_c = VAR_DECL;
tree block_a = BLOCK;
tree block_b = BLOCK;
tree block_c = BLOCK;
BLOCK_VARS(block_a) = decl_a;
BLOCK_SUBBLOCKS(block_a) = block_b;
BLOCK_CHAIN(block_a) = block_c;
BLOCK_SUPERCONTEXT(block_a) = foo;
BLOCK_VARS(block_b) = decl_b;
BLOCK_SUPERCONTEXT(block_b) = block_a;
BLOCK_VARS(block_c) = decl_c;
BLOCK_SUPERCONTEXT(block_c) = foo;
DECL_INITIAL(foo) = block_a;
```

11.8.2 Function Properties

To determine the scope of a function, you can use the DECL_CONTEXT macro. This macro will return the class (either a RECORD_TYPE or a UNION_TYPE) or namespace (a NAMESPACE_DECL) of which the function is a member. For a virtual function, this macro returns the class in which the function was actually defined, not the base class in which the virtual declaration occurred.
In C, the `DECL_CONTEXT` for a function maybe another function. This representation indicates that the GNU nested function extension is in use. For details on the semantics of nested functions, see the GCC Manual. The nested function can refer to local variables in its containing function. Such references are not explicitly marked in the tree structure; back ends must look at the `DECL_CONTEXT` for the referenced `VAR_DECL`. If the `DECL_CONTEXT` for the referenced `VAR_DECL` is not the same as the function currently being processed, and neither `DECL_EXTERNAL` nor `TREE_STATIC` hold, then the reference is to a local variable in a containing function, and the back end must take appropriate action.

`DECL_EXTERNAL`
This predicate holds if the function is undefined.

`TREE_PUBLIC`
This predicate holds if the function has external linkage.

`TREE_STATIC`
This predicate holds if the function has been defined.

`TREE_THIS_VOLATILE`
This predicate holds if the function does not return normally.

`TREE_READONLY`
This predicate holds if the function can only read its arguments.

`DECL_PURE_P`
This predicate holds if the function can only read its arguments, but may also read global memory.

`DECL_VIRTUAL_P`
This predicate holds if the function is virtual.

`DECL_ARTIFICIAL`
This macro holds if the function was implicitly generated by the compiler, rather than explicitly declared. In addition to implicitly generated class member functions, this macro holds for the special functions created to implement static initialization and destruction, to compute run-time type information, and so forth.

`DECL_FUNCTION_SPECIFIC_TARGET`
This macro returns a tree node that holds the target options that are to be used to compile this particular function or `NULL_TREE` if the function is to be compiled with the target options specified on the command line.

`DECL_FUNCTION_SPECIFIC_OPTIMIZATION`
This macro returns a tree node that holds the optimization options that are to be used to compile this particular function or `NULL_TREE` if the function is to be compiled with the optimization options specified on the command line.

11.9 Language-dependent trees
Front ends may wish to keep some state associated with various GENERIC trees while parsing. To support this, trees provide a set of flags that may be used by the front end. They are accessed using `TREE_LANG_FLAG_n` where ‘n’ is currently 0 through 6.
If necessary, a front end can use some language-dependent tree codes in its GENERIC representation, so long as it provides a hook for converting them to GIMPLE and doesn’t expect them to work with any (hypothetical) optimizers that run before the conversion to GIMPLE. The intermediate representation used while parsing C and C++ looks very little like GENERIC, but the C and C++ gimplifier hooks are perfectly happy to take it as input and spit out GIMPLE.

11.10 C and C++ Trees

This section documents the internal representation used by GCC to represent C and C++ source programs. When presented with a C or C++ source program, GCC parses the program, performs semantic analysis (including the generation of error messages), and then produces the internal representation described here. This representation contains a complete representation for the entire translation unit provided as input to the front end. This representation is then typically processed by a code-generator in order to produce machine code, but could also be used in the creation of source browsers, intelligent editors, automatic documentation generators, interpreters, and any other programs needing the ability to process C or C++ code.

This section explains the internal representation. In particular, it documents the internal representation for C and C++ source constructs, and the macros, functions, and variables that can be used to access these constructs. The C++ representation is largely a superset of the representation used in the C front end. There is only one construct used in C that does not appear in the C++ front end and that is the GNU “nested function” extension. Many of the macros documented here do not apply in C because the corresponding language constructs do not appear in C.

The C and C++ front ends generate a mix of GENERIC trees and ones specific to C and C++. These language-specific trees are higher-level constructs than the ones in GENERIC to make the parser’s job easier. This section describes those trees that aren’t part of GENERIC as well as aspects of GENERIC trees that are treated in a language-specific manner.

If you are developing a “back end”, be it is a code-generator or some other tool, that uses this representation, you may occasionally find that you need to ask questions not easily answered by the functions and macros available here. If that situation occurs, it is quite likely that GCC already supports the functionality you desire, but that the interface is simply not documented here. In that case, you should ask the GCC maintainers (via mail to gcc@gcc.gnu.org) about documenting the functionality you require. Similarly, if you find yourself writing functions that do not deal directly with your back end, but instead might be useful to other people using the GCC front end, you should submit your patches for inclusion in GCC.

11.10.1 Types for C++

In C++, an array type is not qualified; rather the type of the array elements is qualified. This situation is reflected in the intermediate representation. The macros described here will always examine the qualification of the underlying element type when applied to an array type. (If the element type is itself an array, then the recursion continues until a non-array type is found, and the qualification of this type is examined.) So, for example, CP_TYPE_CONST_P will hold of the type const int () [7], denoting an array of seven ints.
The following functions and macros deal with cv-qualification of types:

**CP_TYPE_QUALS**
This macro returns the set of type qualifiers applied to this type. This value is **TYPE_UNQUALIFIED** if no qualifiers have been applied. The **TYPEQUAL_CONST** bit is set if the type is **const**-qualified. The **TYPEQUAL_VOLATILE** bit is set if the type is **volatile**-qualified. The **TYPEQUAL_RESTRICT** bit is set if the type is **restrict**-qualified.

**CP_TYPE_CONST_P**
This macro holds if the type is **const**-qualified.

**CP_TYPE_VOLATILE_P**
This macro holds if the type is **volatile**-qualified.

**CP_TYPE_RESTRICT_P**
This macro holds if the type is **restrict**-qualified.

**CP_TYPE_CONST_NON_VOLATILE_P**
This predicate holds for a type that is **const**-qualified, but not **volatile**-qualified; other cv-qualifiers are ignored as well: only the **const**-ness is tested.

A few other macros and functions are usable with all types:

**TYPE_SIZE**
The number of bits required to represent the type, represented as an **INTEGER_CST**. For an incomplete type, **TYPE_SIZE** will be **NULL_TREE**.

**TYPE_ALIGN**
The alignment of the type, in bits, represented as an **int**.

**TYPE_NAME**
This macro returns a declaration (in the form of a **TYPE_DECL**) for the type. (Note this macro does not return an **IDENTIFIER_NODE**, as you might expect, given its name!) You can look at the **DECL_NAME** of the **TYPE_DECL** to obtain the actual name of the type. The **TYPE_NAME** will be **NULL_TREE** for a type that is not a built-in type, the result of a typedef, or a named class type.

**CP_INTEGRAL_TYPE**
This predicate holds if the type is an integral type. Notice that in C++, enumerations are not integral types.

**ARITHMETIC_TYPE_P**
This predicate holds if the type is an integral type (in the C++ sense) or a floating point type.

**CLASS_TYPE_P**
This predicate holds for a class-type.

**TYPE_BUILT_IN**
This predicate holds for a built-in type.

**TYPE_PTRMEM_P**
This predicate holds if the type is a pointer to data member.
TYPE_PTR_P
This predicate holds if the type is a pointer type, and the pointee is not a data
member.

TYPE_PTRFN_P
This predicate holds for a pointer to function type.

TYPE_PTROB_P
This predicate holds for a pointer to object type. Note however that it does not
hold for the generic pointer to object type void *. You may use TYPE_PTROBV_P
to test for a pointer to object type as well as void *.

The table below describes types specific to C and C++ as well as language-dependent info
about GENERIC types.

POINTER_TYPE
Used to represent pointer types, and pointer to data member types. If TREE_TYPE
is a pointer to data member type, then TYPE_PTRMEM_P will hold. For a
pointer to data member type of the form ’T X::*’, TYPE_PTRMEM_CLASS_TYPE
will be the type X, while TYPE_PTRMEM_POINTED_TO_TYPE will be the type T.

RECORD_TYPE
Used to represent struct and class types in C and C++. If TYPE_PTRMEMFUNC_P
holds, then this type is a pointer-to-member type. In that case, the TYPE_PTRMEMFUNC_FN_TYPE
is a POINTER_TYPE pointing to a METHOD_TYPE. The
METHOD_TYPE is the type of a function pointed to by the pointer-to-member
function. If TYPE_PTRMEMFUNC_P does not hold, this type is a class type. For
more information, see Section 11.10.3 [Classes], page 212.

UNKNOWN_TYPE
This node is used to represent a type the knowledge of which is insufficient for
a sound processing.

TYPENAME_TYPE
Used to represent a construct of the form typename T::A. The TYPE_CONTEXT
is T; the TYPE_NAME is an IDENTIFIER_NODE for A. If the type is specified via a
template-id, then TYPENAME_TYPE_FULLNAME yields a TEMPLATE_ID_EXPR. The
TREE_TYPE is non-NULL if the node is implicitly generated in support for the
implicit typename extension; in which case the TREE_TYPE is a type node for
the base-class.

TYPEOF_TYPE
Used to represent the __typeof__ extension. The TYPE_FIELDS is the expres-
sion the type of which is being represented.

11.10.2 Namespaces
The root of the entire intermediate representation is the variable global_namespace. This is
the namespace specified with :: in C++ source code. All other namespaces, types, variables,
functions, and so forth can be found starting with this namespace.

However, except for the fact that it is distinguished as the root of the representation,
the global namespace is no different from any other namespace. Thus, in what follows, we
describe namespaces generally, rather than the global namespace in particular.
A namespace is represented by a NAMESPACE_DECL node.

The following macros and functions can be used on a NAMESPACE_DECL:

**DECL_NAME**
This macro is used to obtain the IDENTIFIER_NODE corresponding to the un-
qualified name of the name of the namespace (see Section 11.2.2 [Identifiers],
page 179). The name of the global namespace is ‘::’, even though in C++
the global namespace is unnamed. However, you should use comparison with
global_namespace, rather than DECL_NAME to determine whether or not a
namespace is the global one. An unnamed namespace will have a DECL_NAME
equal to anonymous_namespace_name. Within a single translation unit, all un-
named namespaces will have the same name.

**DECL_CONTEXT**
This macro returns the enclosing namespace. The DECL_CONTEXT for the
global_namespace is NULL_TREE.

**DECL_NAMESPACE_ALIAS**
If this declaration is for a namespace alias, then DECL_NAMESPACE_ALIAS is the
namespace for which this one is an alias.

Do not attempt to use cp_namespace_decls for a namespace which is an alias.
Instead, follow DECL_NAMESPACE_ALIAS links until you reach an ordinary, non-
alias, namespace, and call cp_namespace_decls there.

**DECL_NAMESPACE_STD_P**
This predicate holds if the namespace is the special ::std namespace.

**cp_namespace_decls**
This function will return the declarations contained in the namespace, including
types, overloaded functions, other namespaces, and so forth. If there are no decl-
arations, this function will return NULL_TREE. The declarations are connected
through their TREE_CHAIN fields.

Although most entries on this list will be declarations, TREE_LIST nodes may
also appear. In this case, the TREE_VALUE will be an OVERLOAD. The value of the
TREE_PURPOSE is unspecified; back ends should ignore this value. As with the
other kinds of declarations returned by cp_namespace_decls, the TREE_CHAIN
will point to the next declaration in this list.

For more information on the kinds of declarations that can occur on this list,
See Section 11.4 [Declarations], page 184. Some declarations will not appear on
this list. In particular, no FIELDDECL, LABELDECL, or PARMDECL nodes will
appear here.

This function cannot be used with namespaces that have DECL_NAMESPACE_-
ALIAS set.

### 11.10.3 Classes

Besides namespaces, the other high-level scoping construct in C++ is the class. (Through-
out this manual the term class is used to mean the types referred to in the ANSI/ISO
C++ Standard as classes; these include types defined with the class, struct, and union
keywords.)
A class type is represented by either a RECORD_TYPE or a UNION_TYPE. A class declared with the union tag is represented by a UNION_TYPE, while classes declared with either the struct or the class tag are represented by RECORD_TYPES. You can use the CLASSTYPE_DECLARED_CLASS macro to discern whether or not a particular type is a class as opposed to a struct. This macro will be true only for classes declared with the class tag.

Almost all non-function members are available on the TYPE_FIELDS list. Given one member, the next can be found by following the TREE_CHAIN. You should not depend in any way on the order in which fields appear on this list. All nodes on this list will be ‘DECL’ nodes. A FIELDDECL is used to represent a non-static data member, a VARDECL is used to represent a static data member, and a TYPEDECL is used to represent a type. Note that the CONSTDECL for an enumeration constant will appear on this list, if the enumeration type was declared in the class. (Of course, the TYPEDECL for the enumeration type will appear here as well.) There are no entries for base classes on this list. In particular, there is no FIELDDECL for the “base-class portion” of an object.

The TYPE_VFIELD is a compiler-generated field used to point to virtual function tables. It may or may not appear on the TYPE_FIELDS list. However, back ends should handle the TYPE_VFIELD just like all the entries on the TYPE_FIELDS list.

The function members are available on the TYPE_METHODS list. Again, subsequent members are found by following the TREE_CHAIN field. If a function is overloaded, each of the overloaded functions appears; no OVERLOAD nodes appear on the TYPE_METHODS list. Implicitly declared functions (including default constructors, copy constructors, assignment operators, and destructors) will appear on this list as well.

Every class has an associated binfo, which can be obtained with TYPE_BINFO. Binfos are used to represent base-classes. The binfo given by TYPE_BINFO is the degenerate case, whereby every class is considered to be its own base-class. The base binfos for a particular binfo are held in a vector, whose length is obtained with BINFO_N_BASE_BINFOS. The base binfos themselves are obtained with BINFO_BASE_BINFO and BINFO_BASE_ITERATE. To add a new binfo, use BINFO_BASE_APPEND. The vector of base binfos can be obtained with BINFO_BASE_BINFOS, but normally you do not need to use that. The class type associated with a binfo is given by BINFO_TYPE. It is not always the case that BINFO_TYPE (TYPE_BINFO (x)) is the same binfo as y. The reason is that if y is a binfo representing a base-class B of a derived class D, then BINFO_TYPE (y) will be B, and TYPE_BINFO (BINFO_TYPE (y)) will be B as its own base-class, rather than as a base-class of D.

The access to a base type can be found with BINFO_BASE_ACCESS. This will produce access_public_node, access_private_node or access_protected_node. If bases are always public, BINFO_BASE_ACCESS may be NULL.

BINFO_VIRTUAL_P is used to specify whether the binfo is inherited virtually or not. The other flags, BINFO_MARKED_P and BINFO_FLAG_1 to BINFO_FLAG_6 can be used for language specific use.

The following macros can be used on a tree node representing a class-type.

LOCAL_CLASS_P
This predicate holds if the class is local class i.e. declared inside a function body.
TYPE_POLYMORPHIC_P
This predicate holds if the class has at least one virtual function (declared or inherited).

TYPE_HAS_DEFAULT_CONSTRUCTOR
This predicate holds whenever its argument represents a class-type with default constructor.

CLASSTYPE_HAS_MUTABLE
TYPE_HAS_MUTABLE_P
These predicates hold for a class-type having a mutable data member.

CLASSTYPE_NON_POD_P
This predicate holds only for class-types that are not PODs.

TYPE_HAS_NEW_OPERATOR
This predicate holds for a class-type that defines operator new.

TYPE_HAS_ARRAY_NEW_OPERATOR
This predicate holds for a class-type for which operator new[] is defined.

TYPE_OVERLOADS_CALL_EXPR
This predicate holds for class-type for which the function call operator() is overloaded.

TYPE_OVERLOADS_ARRAY_REF
This predicate holds for a class-type that overloads operator[].

TYPE_OVERLOADS_ARROW
This predicate holds for a class-type for which operator-> is overloaded.

11.10.4 Functions for C++
A function is represented by a FUNCTION_DECL node. A set of overloaded functions is sometimes represented by an OVERLOAD node.

An OVERLOAD node is not a declaration, so none of the ‘DECL_’ macros should be used on an OVERLOAD. An OVERLOAD node is similar to a TREE_LIST. Use OVL_CURRENT to get the function associated with an OVERLOAD node; use OVL_NEXT to get the next OVERLOAD node in the list of overloaded functions. The macros OVL_CURRENT and OVL_NEXT are actually polymorphic; you can use them to work with FUNCTION_DECL nodes as well as with overloads. In the case of a FUNCTION_DECL, OVL_CURRENT will always return the function itself, and OVL_NEXT will always be NULL_TREE.

To determine the scope of a function, you can use the DECL_CONTEXT macro. This macro will return the class (either a RECORD_TYPE or a UNION_TYPE) or namespace (a NAMESPACE_DECL) of which the function is a member. For a virtual function, this macro returns the class in which the function was actually defined, not the base class in which the virtual declaration occurred.

If a friend function is defined in a class scope, the DECL_FRIEND_CONTEXT macro can be used to determine the class in which it was defined. For example, in

```cpp
class C { friend void f() {} ; }
```

the DECL_CONTEXT for f will be the global_namespace, but the DECL_FRIEND_CONTEXT will be the RECORD_TYPE for C.
The following macros and functions can be used on a `FUNCTION_DECL`:

**DECL_MAIN_P**
This predicate holds for a function that is the program entry point.

**DECL_LOCAL_FUNCTION_P**
This predicate holds if the function was declared at block scope, even though it has a global scope.

**DECL_ANTICIPATED**
This predicate holds if the function is a built-in function but its prototype is not yet explicitly declared.

**DECL_EXTERN_C_FUNCTION_P**
This predicate holds if the function is declared as an `extern "C"` function.

**DECL_LINKONCE_P**
This macro holds if multiple copies of this function may be emitted in various translation units. It is the responsibility of the linker to merge the various copies. Template instantiations are the most common example of functions for which `DECL_LINKONCE_P` holds; G++ instantiates needed templates in all translation units which require them, and then relies on the linker to remove duplicate instantiations.

FIXME: This macro is not yet implemented.

**DECL_FUNCTION_MEMBER_P**
This macro holds if the function is a member of a class, rather than a member of a namespace.

**DECL_STATIC_FUNCTION_P**
This predicate holds if the function a static member function.

**DECL_NONSTATIC_MEMBER_FUNCTION_P**
This macro holds for a non-static member function.

**DECL_CONST_MEMFUNC_P**
This predicate holds for a `const`-member function.

**DECL_VOLATILE_MEMFUNC_P**
This predicate holds for a `volatile`-member function.

**DECL_CONSTRUCTOR_P**
This macro holds if the function is a constructor.

**DECL_NONCONVERTING_P**
This predicate holds if the constructor is a non-converting constructor.

**DECL_COMPLETE_CONSTRUCTOR_P**
This predicate holds for a function which is a constructor for an object of a complete type.

**DECL_BASE_CONSTRUCTOR_P**
This predicate holds for a function which is a constructor for a base class sub-object.
 DECL_COPY_CONSTRUCTOR_P
This predicate holds for a function which is a copy-constructor.

DECL_DESTRUCTOR_P
This macro holds if the function is a destructor.

DECL_COMPLETE_DESTRUCTOR_P
This predicate holds if the function is the destructor for an object a complete
type.

DECL_OVERLOADED_OPERATOR_P
This macro holds if the function is an overloaded operator.

DECL_CONV_FN_P
This macro holds if the function is a type-conversion operator.

DECL_GLOBAL_CTOR_P
This predicate holds if the function is a file-scope initialization function.

DECL_GLOBAL_DTOR_P
This predicate holds if the function is a file-scope finalization function.

DECL_THUNK_P
This predicate holds if the function is a thunk.
These functions represent stub code that adjusts the this pointer and then
jumps to another function. When the jumped-to function returns, control is
transferred directly to the caller, without returning to the thunk. The first
parameter to the thunk is always the this pointer; the thunk should add THUNK_
DELTA to this value. (The THUNK_DELTA is an int, not an INTEGER_CST.)

Then, if THUNK_VCALL_OFFSET (an INTEGER_CST) is nonzero the adjusted this
pointer must be adjusted again. The complete calculation is given by the fol-
lowing pseudo-code:

\[
\text{this} += \text{THUNK\_DELTA} \\
\text{if} (\text{THUNK\_VCALL\_OFFSET}) \text{ this} += (*((ptrdiff_t **) this)[\text{THUNK\_VCALL\_OFFSET}]
\]

Finally, the thunk should jump to the location given by DECL_INITIAL; this
will always be an expression for the address of a function.

DECL_NONThunk_FUNCTION_P
This predicate holds if the function is not a thunk function.

GLOBAL_INIT_PRIORITY
If either DECL_GLOBAL_CTOR_P or DECL_GLOBAL_DTOR_P holds, then this gives
the initialization priority for the function. The linker will arrange that all
functions for which DECL_GLOBAL_CTOR_P holds are run in increasing order of
priority before main is called. When the program exits, all functions for which
DECL_GLOBAL_DTOR_P holds are run in the reverse order.

TYPE_RAISES_EXCEPTIONS
This macro returns the list of exceptions that a (member-)function can raise.
The returned list, if non NULL, is comprised of nodes whose TREE_VALUE repre-
sents a type.
TYPE_NOTHROW_P

This predicate holds when the exception-specification of its arguments is of the form `()`.

DECL_ARRAY_DELETE_OPERATOR_P

This predicate holds if the function an overloaded `operator delete[]`

11.10.5 Statements for C++

A function that has a definition in the current translation unit will have a non-NULL DECL_INITIAL. However, back ends should not make use of the particular value given by DECL_INITIAL.

The DECL_SAVED_TREE macro will give the complete body of the function.

11.10.5.1 Statements

There are tree nodes corresponding to all of the source-level statement constructs, used within the C and C++ frontends. These are enumerated here, together with a list of the various macros that can be used to obtain information about them. There are a few macros that can be used with all statements:

STMT_IS_FULL_EXPR_P

In C++, statements normally constitute “full expressions”; temporaries created during a statement are destroyed when the statement is complete. However, G++ sometimes represents expressions by statements; these statements will not have STMT_IS_FULL_EXPR_P set. Temporaries created during such statements should be destroyed when the innermost enclosing statement with STMT_IS_FULL_EXPR_P set is exited.

Here is the list of the various statement nodes, and the macros used to access them. This documentation describes the use of these nodes in non-template functions (including instantiations of template functions). In template functions, the same nodes are used, but sometimes in slightly different ways.

Many of the statements have substatements. For example, a `while` loop will have a body, which is itself a statement. If the substatement is NULL_TREE, it is considered equivalent to a statement consisting of a single `;`, i.e., an expression statement in which the expression has been omitted. A substatement may in fact be a list of statements, connected via their TREE_CHAINs. So, you should always process the statement tree by looping over substatements, like this:

```c
void process_stmt (stmt)
  tree stmt;
{
  while (stmt)
  {
    switch (TREE_CODE (stmt))
    {
    case IF_STMT:
      process_stmt (THEN_CLAUSE (stmt));
      /* More processing here. */
      break;
    ...
```
In other words, while the then clause of an if statement in C++ can be only one statement (although that one statement may be a compound statement), the intermediate representation will sometimes use several statements chained together.

**BREAK_STMT**
Used to represent a `break` statement. There are no additional fields.

**CLEANUP_STMT**
Used to represent an action that should take place upon exit from the enclosing scope. Typically, these actions are calls to destructors for local objects, but back ends cannot rely on this fact. If these nodes are in fact representing such destructors, `CLEANUPDECL` will be the `VARDECL` destroyed. Otherwise, `CLEANUPDECL` will be `NULL_TREE`. In any case, the `CLEANUPEXP` is the expression to execute. The cleanups executed on exit from a scope should be run in the reverse order of the order in which the associated `CLEANUP_STMTs` were encountered.

**CONTINUE_STMT**
Used to represent a `continue` statement. There are no additional fields.

**CTOR_STMT**
Used to mark the beginning (if `CTORBEGIN_P` holds) or end (if `CTOREND_P` holds) of the main body of a constructor. See also `SUBOBJECT` for more information on how to use these nodes.

**DO_STMT**
Used to represent a `do` loop. The body of the loop is given by `DO_BODY` while the termination condition for the loop is given by `DO_COND`. The condition for a `do`-statement is always an expression.

**EMPTY_CLASS_EXPR**
Used to represent a temporary object of a class with no data whose address is never taken. (All such objects are interchangeable.) The `TREE_TYPE` represents the type of the object.

**EXPR_STMT**
Used to represent an expression statement. Use `EXPR_STMT_EXPR` to obtain the expression.

**FOR_STMT**
Used to represent a `for` statement. The `FOR_INIT_STMT` is the initialization statement for the loop. The `FOR_COND` is the termination condition. The `FOR_EXPR` is the expression executed right before the `FOR_COND` on each loop iteration; often, this expression increments a counter. The body of the loop is given by `FOR_BODY`. Note that `FOR_INIT_STMT` and `FOR_BODY` return statements, while `FOR_COND` and `FOR_EXPR` return expressions.
**HANDLER**

Used to represent a C++ catch block. The `HANDLER_TYPE` is the type of exception that will be caught by this handler; it is equal (by pointer equality) to `NULL` if this handler is for all types. `HANDLER_PARMS` is the `DECL_STMT` for the catch parameter, and `HANDLER_BODY` is the code for the block itself.

**IF_STMT**

Used to represent an if statement. The `IF_COND` is the expression. If the condition is a `TREE_LIST`, then the `TREE_PURPOSE` is a statement (usually a `DECL_STMT`). Each time the condition is evaluated, the statement should be executed. Then, the `TREE_VALUE` should be used as the conditional expression itself. This representation is used to handle C++ code like this:

```cpp
if (int i = 7) ...
```

where there is a new local variable (or variables) declared within the condition. The `THEN_CLAUSE` represents the statement given by the `then` condition, while the `ELSE_CLAUSE` represents the statement given by the `else` condition.

**SUBOBJECT**

In a constructor, these nodes are used to mark the point at which a subobject of `this` is fully constructed. If, after this point, an exception is thrown before a `CTOR_STMT` with `CTOR_END_P` set is encountered, the `SUBOBJECT_CLEANUP` must be executed. The cleanups must be executed in the reverse order in which they appear.

**SWITCH_STMT**

Used to represent a switch statement. The `SWITCH_STMT_COND` is the expression on which the switch is occurring. See the documentation for an `IF_STMT` for more information on the representation used for the condition. The `SWITCH_STMT_BODY` is the body of the switch statement. The `SWITCH_STMT_TYPE` is the original type of switch expression as given in the source, before any compiler conversions.

**TRY_BLOCK**

Used to represent a try block. The body of the try block is given by `TRY_STMTS`. Each of the catch blocks is a `HANDLER` node. The first handler is given by `TRY_HANDLERS`. Subsequent handlers are obtained by following the `TREE_CHAIN` link from one handler to the next. The body of the handler is given by `HANDLER_BODY`.

If `CLEANUP_P` holds of the `TRY_BLOCK`, then the `TRY_HANDLERS` will not be a `HANDLER` node. Instead, it will be an expression that should be executed if an exception is thrown in the try block. It must rethrow the exception after executing that code. And, if an exception is thrown while the expression is executing, `terminate` must be called.

**USING_STMT**

Used to represent a using directive. The namespace is given by `USING_STMT_NAMESPACE`, which will be a `NAMESPACE_DECLS`. This node is needed inside template functions, to implement using directives during instantiation.
WHILE_STMT

Used to represent a `while` loop. The `WHILE_COND` is the termination condition for the loop. See the documentation for an `IF_STMT` for more information on the representation used for the condition.

The `WHILE_BODY` is the body of the loop.

### 11.10.6 C++ Expressions

This section describes expressions specific to the C and C++ front ends.

**TYPEID_EXPR**

Used to represent a `typeid` expression.

**NEW_EXPR**

**VEC_NEW_EXPR**

Used to represent a call to `new` and `new[]` respectively.

**DELETE_EXPR**

**VEC_DELETE_EXPR**

Used to represent a call to `delete` and `delete[]` respectively.

**MEMBER_REF**

Represents a reference to a member of a class.

**THROW_EXPR**

Represents an instance of `throw` in the program. Operand 0, which is the expression to throw, may be `NULL_TREE`.

**AGGR_INIT_EXPR**

An `AGGR_INIT_EXPR` represents the initialization as the return value of a function call, or as the result of a constructor. An `AGGR_INIT_EXPR` will only appear as a full-expression, or as the second operand of a `TARGET_EXPR`. `AGGR_INIT_EXPRs` have a representation similar to that of `CALL_EXPRs`. You can use the `AGGR_INIT_EXPR_FN` and `AGGR_INIT_EXPR_ARG` macros to access the function to call and the arguments to pass.

If `AGGR_INIT_VIA_CTOR_P` holds of the `AGGR_INIT_EXPR`, then the initialization is via a constructor call. The address of the `AGGR_INIT_EXPR_SLOT` operand, which is always a `VAR_DECL`, is taken, and this value replaces the first argument in the argument list.

In either case, the expression is void.

### 11.11 Java Trees
12 GIMPLE

GIMPLE is a three-address representation derived from GENERIC by breaking down GENERIC expressions into tuples of no more than 3 operands (with some exceptions like function calls). GIMPLE was heavily influenced by the SIMPLE IL used by the McCAT compiler project at McGill University, though we have made some different choices. For one thing, SIMPLE doesn’t support goto.

Temporaries are introduced to hold intermediate values needed to compute complex expressions. Additionally, all the control structures used in GENERIC are lowered into conditional jumps, lexical scopes are removed and exception regions are converted into an on the side exception region tree.

The compiler pass which converts GENERIC into GIMPLE is referred to as the ‘gimplifier’. The gimplifier works recursively, generating GIMPLE tuples out of the original GENERIC expressions.

One of the early implementation strategies used for the GIMPLE representation was to use the same internal data structures used by front ends to represent parse trees. This simplified implementation because we could leverage existing functionality and interfaces. However, GIMPLE is a much more restrictive representation than abstract syntax trees (AST), therefore it does not require the full structural complexity provided by the main tree data structure.

The GENERIC representation of a function is stored in the DECL_SAVED_TREE field of the associated FUNCTION_DECL tree node. It is converted to GIMPLE by a call to gimplify_function_tree.

If a front end wants to include language-specific tree codes in the tree representation which it provides to the back end, it must provide a definition of LANG_HOOKS_GIMPLIFY_EXPR which knows how to convert the front end trees to GIMPLE. Usually such a hook will involve much of the same code for expanding front end trees to RTL. This function can return fully lowered GIMPLE, or it can return GENERIC trees and let the main gimplifier lower them the rest of the way; this is often simpler. GIMPLE that is not fully lowered is known as ‘High GIMPLE’ and consists of the IL before the pass pass_lower_cf. High GIMPLE contains some container statements like lexical scopes (represented by GIMPLE_BIND) and nested expressions (e.g., GIMPLE_TRY), while “Low GIMPLE” exposes all of the implicit jumps for control and exception expressions directly in the IL and EH region trees.

The C and C++ front ends currently convert directly from front end trees to GIMPLE, and hand that off to the back end rather than first converting to GENERIC. Their gimplifier hooks know about all the _STMT nodes and how to convert them to GENERIC forms. There was some work done on a genericization pass which would run first, but the existence of STMT_EXPR meant that in order to convert all of the C statements into GENERIC equivalents would involve walking the entire tree anyway, so it was simpler to lower all the way. This might change in the future if someone writes an optimization pass which would work better with higher-level trees, but currently the optimizers all expect GIMPLE.

You can request to dump a C-like representation of the GIMPLE form with the flag ‘-fdump-tree-gimple’.
12.1 Tuple representation

GIMPLE instructions are tuples of variable size divided in two groups: a header describing the instruction and its locations, and a variable length body with all the operands. Tuples are organized into a hierarchy with 3 main classes of tuples.

12.1.1 gimple_statement_base (gsbase)

This is the root of the hierarchy, it holds basic information needed by most GIMPLE statements. There are some fields that may not be relevant to every GIMPLE statement, but those were moved into the base structure to take advantage of holes left by other fields (thus making the structure more compact). The structure takes 4 words (32 bytes) on 64 bit hosts:

```
Field                Size (bits)
code                 8
subcode              16
no_warning           1
visited              1
nontemporal_move     1
plf                  2
modified             1
has_volatile_ops     1
references_memory_p  1
uid                  32
location             32
num_ops              32
bb                   64
block                63
```

Total size 32 bytes

- **code**: Main identifier for a GIMPLE instruction.
- **subcode**: Used to distinguish different variants of the same basic instruction or provide flags applicable to a given code. The `subcode` flags field has different uses depending on the code of the instruction, but mostly it distinguishes instructions of the same family. The most prominent use of this field is in assignments, where subcode indicates the operation done on the RHS of the assignment. For example, `a = b + c` is encoded as `GIMPLE_ASSIGN <PLUS_EXPR, a, b, c>`.
- **no_warning**: Bitflag to indicate whether a warning has already been issued on this statement.
- **visited**: General purpose “visited” marker. Set and cleared by each pass when needed.
- **nontemporal_move**: Bitflag used in assignments that represent non-temporal moves. Although this bitflag is only used in assignments, it was moved into the base to take advantage of the bit holes left by the previous fields.
- **plf**: Pass Local Flags. This 2-bit mask can be used as general purpose markers by any pass. Passes are responsible for clearing and setting these two flags accordingly.
- **modified**: Bitflag to indicate whether the statement has been modified. Used mainly by the operand scanner to determine when to re-scan a statement for operands.
• **has_volatile_ops** Bitflag to indicate whether this statement contains operands that have been marked volatile.

• **references_memory_p** Bitflag to indicate whether this statement contains memory references (i.e., its operands are either global variables, or pointer dereferences or anything that must reside in memory).

• **uid** This is an unsigned integer used by passes that want to assign IDs to every statement. These IDs must be assigned and used by each pass.

• **location** This is a `location_t` identifier to specify source code location for this statement. It is inherited from the front end.

• **num_ops** Number of operands that this statement has. This specifies the size of the operand vector embedded in the tuple. Only used in some tuples, but it is declared in the base tuple to take advantage of the 32-bit hole left by the previous fields.

• **bb** Basic block holding the instruction.

• **block** Lexical block holding this statement. Also used for debug information generation.

12.1.2 gimple_statement_with_ops

This tuple is actually split in two: `gimple_statement_with_ops_base` and `gimple_statement_with_ops`. This is needed to accommodate the way the operand vector is allocated. The operand vector is defined to be an array of 1 element. So, to allocate a dynamic number of operands, the memory allocator (`gimple_alloc`) simply allocates enough memory to hold the structure itself plus \( N - 1 \) operands which run “off the end” of the structure. For example, to allocate space for a tuple with 3 operands, `gimple_alloc` reserves `sizeof (struct gimple_statement_with_ops) + 2 * sizeof (tree)` bytes.

On the other hand, several fields in this tuple need to be shared with the `gimple_statement_with_memory_ops` tuple. So, these common fields are placed in `gimple_statement_with_ops_base` which is then inherited from the other two tuples.

```
gsbase 256
def_ops 64
use_ops 64
op num_ops * 64
Total 48 + 8 * num_ops bytes
```

• **gsbase** Inherited from `struct gimple_statement_base`.

• **def_ops** Array of pointers into the operand array indicating all the slots that contain a variable written-to by the statement. This array is also used for immediate use chaining. Note that it would be possible to not rely on this array, but the changes required to implement this are pretty invasive.

• **use_ops** Similar to **def_ops** but for variables read by the statement.

• **op** Array of trees with `num_ops` slots.

12.1.3 gimple_statement_with_memory_ops

This tuple is essentially identical to `gimple_statement_with_ops`, except that it contains 4 additional fields to hold vectors related memory stores and loads. Similar to the pre-
previous case, the structure is split in two to accommodate for the operand vector (**gimple_statement_with_memory_ops_base** and **gimple_statement_with_memory_ops**).

<table>
<thead>
<tr>
<th>Field</th>
<th>Size (bits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>gsbase</td>
<td>256</td>
</tr>
<tr>
<td>def_ops</td>
<td>64</td>
</tr>
<tr>
<td>use_ops</td>
<td>64</td>
</tr>
<tr>
<td>vdef_ops</td>
<td>64</td>
</tr>
<tr>
<td>vuse_ops</td>
<td>64</td>
</tr>
<tr>
<td>stores</td>
<td>64</td>
</tr>
<tr>
<td>loads</td>
<td>64</td>
</tr>
<tr>
<td>op</td>
<td>num_ops * 64</td>
</tr>
</tbody>
</table>

Total size: $80 + 8 \times \text{num_ops}$ bytes

- **vdef_ops** Similar to **def_ops** but for VDEF operators. There is one entry per memory symbol written by this statement. This is used to maintain the memory SSA use-def and def-def chains.

- **vuse_ops** Similar to **use_ops** but for VUSE operators. There is one entry per memory symbol loaded by this statement. This is used to maintain the memory SSA use-def chains.

- **stores** Bitset with all the UIDs for the symbols written-to by the statement. This is different than **vdef_ops** in that all the affected symbols are mentioned in this set. If memory partitioning is enabled, the **vdef_ops** vector will refer to memory partitions. Furthermore, no SSA information is stored in this set.

- **loads** Similar to **stores**, but for memory loads. (Note that there is some amount of redundancy here, it should be possible to reduce memory utilization further by removing these sets).

All the other tuples are defined in terms of these three basic ones. Each tuple will add some fields. The main gimple type is defined to be the union of all these structures (**GTY** markers elided for clarity):

```c
union gimple_statement_d {
  struct gimple_statement_base gsbase;
  struct gimple_statement_with_ops gsops;
  struct gimple_statement_with_memory_ops gsmem;
  struct gimple_statement_omp omp;
  struct gimple_statement_bind gimple_bind;
  struct gimple_statement_eh_filter gimple_eh_filter;
  struct gimple_statement_phi gimple_phi;
  struct gimple_statement_resx gimple_resx;
  struct gimple_statement_try gimple_try;
  struct gimple_statement_wce gimple_wce;
  struct gimple_statement_asm gimple_asm;
  struct gimple_statement_omp_critical gimple_omp_critical;
  struct gimple_statement_omp_for gimple_omp_for;
  struct gimple_statement_omp_parallel gimple_omp_parallel;
  struct gimple_statement_omp_task gimple_omp_task;
  struct gimple_statement_omp_sections gimple_omp_sections;
  struct gimple_statement_omp_single gimple_omp_single;
  struct gimple_statement_omp_continue gimple_omp_continue;
  struct gimple_statement_omp_atomic_load gimple_omp_atomic_load;
```

struct gimple_statement_omp_atomic_store gimple_omp_atomic_store;

12.2 GIMPLE instruction set

The following table briefly describes the GIMPLE instruction set.

<table>
<thead>
<tr>
<th>Instruction</th>
<th>High GIMPLE</th>
<th>Low GIMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>GIMPLE_ASM</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_ASSIGN</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_BIND</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_CALL</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_CATCH</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_COND</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_DEBUG</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_EH_FILTER</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_GOTO</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_LABEL</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_NOP</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_ATOMIC_LOAD</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_ATOMIC_STORE</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_CONTINUE</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_CRITICAL</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_FOR</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_MASTER</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_ORDERED</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_PARALLEL</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_RETURN</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_SECTION</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_SECTIONS</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_SECTIONS_SWITCH</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_OMP_SINGLE</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_PHI</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_RESX</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_RETURN</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_SWITCH</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>GIMPLE_TRY</td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

12.3 Exception Handling

Other exception handling constructs are represented using GIMPLE_TRY_CATCH. GIMPLE_TRY_CATCH has two operands. The first operand is a sequence of statements to execute. If executing these statements does not throw an exception, then the second operand is ignored. Otherwise, if an exception is thrown, then the second operand of the GIMPLE_TRY_CATCH is checked. The second operand may have the following forms:

1. A sequence of statements to execute. When an exception occurs, these statements are executed, and then the exception is rethrown.
2. A sequence of **GIMPLE_CATCH** statements. Each **GIMPLE_CATCH** has a list of applicable exception types and handler code. If the thrown exception matches one of the caught types, the associated handler code is executed. If the handler code falls off the bottom, execution continues after the original **GIMPLE_TRY_CATCH**.

3. A **GIMPLE_EH_FILTER** statement. This has a list of permitted exception types, and code to handle a match failure. If the thrown exception does not match one of the allowed types, the associated match failure code is executed. If the thrown exception does match, it continues unwinding the stack looking for the next handler.

Currently throwing an exception is not directly represented in GIMPLE, since it is implemented by calling a function. At some point in the future we will want to add some way to express that the call will throw an exception of a known type.

Just before running the optimizers, the compiler lowers the high-level EH constructs above into a set of ‘goto’s, magic labels, and EH regions. Continuing to unwind at the end of a cleanup is represented with a **GIMPLE_RESX**.

### 12.4 Temporaries

When gimplification encounters a subexpression that is too complex, it creates a new temporary variable to hold the value of the subexpression, and adds a new statement to initialize it before the current statement. These special temporaries are known as ‘expression temporaries’, and are allocated using `get_formal_tmp_var`. The compiler tries to always evaluate identical expressions into the same temporary, to simplify elimination of redundant calculations.

We can only use expression temporaries when we know that it will not be reevaluated before its value is used, and that it will not be otherwise modified\(^1\). Other temporaries can be allocated using `get_initialized_tmp_var` or `create_tmp_var`.

Currently, an expression like `a = b + 5` is not reduced any further. We tried converting it to something like

```
T1 = b + 5;
a = T1;
```

but this bloated the representation for minimal benefit. However, a variable which must live in memory cannot appear in an expression; its value is explicitly loaded into a temporary first. Similarly, storing the value of an expression to a memory variable goes through a temporary.

### 12.5 Operands

In general, expressions in GIMPLE consist of an operation and the appropriate number of simple operands; these operands must either be a GIMPLE value (`is_gimple_val`), i.e. a constant or a register variable. More complex operands are factored out into temporaries, so that

```
a = b + c + d
```

becomes

---

\(^1\) These restrictions are derived from those in Morgan 4.8.
\[ T1 = b + c; \\
    a = T1 + d; \]

The same rule holds for arguments to a GIMPLE_CALL.

The target of an assignment is usually a variable, but can also be a MEM_REF or a compound lvalue as described below.

12.5.1 Compound Expressions
The left-hand side of a C comma expression is simply moved into a separate statement.

12.5.2 Compound Lvalues
Currently compound lvalues involving array and structure field references are not broken down; an expression like \( a.b[2] = 42 \) is not reduced any further (though complex array subscripts are). This restriction is a workaround for limitations in later optimizers; if we were to convert this to
\[
T1 = &a.b; \\
T1[2] = 42; 
\]
alias analysis would not remember that the reference to \( T1[2] \) came by way of \( a.b \), so it would think that the assignment could alias another member of \( a \); this broke struct-alias-1.c. Future optimizer improvements may make this limitation unnecessary.

12.5.3 Conditional Expressions
A C `?:` expression is converted into an if statement with each branch assigning to the same temporary. So,
\[
    a = b \ ? \ c : d; 
\]
becomes
\[
    if (b == 1) 
        T1 = c; \\
    else 
        T1 = d; \\
    a = T1; 
\]

The GIMPLE level if-conversion pass re-introduces `?:` expression, if appropriate. It is used to vectorize loops with conditions using vector conditional operations.

Note that in GIMPLE, if statements are represented using GIMPLE_COND, as described below.

12.5.4 Logical Operators
Except when they appear in the condition operand of a GIMPLE_COND, logical `and` and `or` operators are simplified as follows: \( a = b \ && \ c \) becomes
\[
    T1 = (bool)b; \\
    if (T1 == true) 
        T1 = (bool)c; \\
    a = T1; 
\]

Note that \( T1 \) in this example cannot be an expression temporary, because it has two different assignments.
12.5.5 Manipulating operands

All gimple operands are of type tree. But only certain types of trees are allowed to be used as operand tuples. Basic validation is controlled by the function get_gimple_rhs_class, which given a tree code, returns an enum with the following values of type enum gimple_rhs_class

- GIMPLE_INVALID_RHS The tree cannot be used as a GIMPLE operand.
- GIMPLE_TERNARY_RHS The tree is a valid GIMPLE ternary operation.
- GIMPLE_BINARY_RHS The tree is a valid GIMPLE binary operation.
- GIMPLE_UNARY_RHS The tree is a valid GIMPLE unary operation.
- GIMPLE_SINGLE_RHS The tree is a single object, that cannot be split into simpler operands (for instance, SSA_NAME, VAR_DECL, COMPONENT_REF, etc).

This operand class also acts as an escape hatch for tree nodes that may be flattened out into the operand vector, but would need more than two slots on the RHS. For instance, a COND_EXPR expression of the form (a op b) ? x : y could be flattened out on the operand vector using 4 slots, but it would also require additional processing to distinguish c = a op b from c = a op b ? x : y. Something similar occurs with ASSERT_EXPR. In time, these special case tree expressions should be flattened into the operand vector.

For tree nodes in the categories GIMPLE_TERNARY_RHS, GIMPLE_BINARY_RHS and GIMPLE_UNARY_RHS, they cannot be stored inside tuples directly. They first need to be split into simpler operands (for instance, SSA_NAME, VAR_DECL, COMPONENT_REF, etc).

12.5.6 Operand vector allocation

The operand vector is stored at the bottom of the three tuple structures that accept operands. This means, that depending on the code of a given statement, its operand vector will be at different offsets from the base of the structure. To access tuple operands use the following accessors

unsigned gimple_num_ops (gimple g)  
Returns the number of operands in statement G.

tree gimple_op (gimple g, unsigned i)  
Returns operand I from statement G.

tree * gimple_ops (gimple g)  
Returns a pointer into the operand vector for statement G. This is computed using an internal table called gimple_ops_offset[]. This table is indexed by the gimple code of G.
When the compiler is built, this table is filled-in using the sizes of the structures used by each statement code defined in gimple.def. Since the operand vector is at the bottom of the structure, for a gimple code C the offset is computed as sizeof (struct-of C) - sizeof (tree).

This mechanism adds one memory indirection to every access when using gimple_op(), if this becomes a bottleneck, a pass can choose to memoize the result from gimple_ops() and use that to access the operands.

12.5.7 Operand validation

When adding a new operand to a gimple statement, the operand will be validated according to what each tuple accepts in its operand vector. These predicates are called by the gimple_name_set...

Each tuple will use one of the following predicates (Note, this list is not exhaustive):

bool is_gimple_val (tree t) [GIMPLE function]
Returns true if t is a "GIMPLE value", which are all the non-addressable stack variables (variables for which is_gimple_reg returns true) and constants (expressions for which is_gimple_min_invariant returns true).

bool is_gimple_addressable (tree t) [GIMPLE function]
Returns true if t is a symbol or memory reference whose address can be taken.

bool is_gimple_asm_val (tree t) [GIMPLE function]
Similar to is_gimple_val but it also accepts hard registers.

bool is_gimple_call_addr (tree t) [GIMPLE function]
Return true if t is a valid expression to use as the function called by a GIMPLE_CALL.

bool is_gimple_mem_ref_addr (tree t) [GIMPLE function]
Return true if t is a valid expression to use as first operand of a MEM_REF expression.

bool is_gimple_constant (tree t) [GIMPLE function]
Return true if t is a valid gimple constant.

bool is_gimple_min_invariant (tree t) [GIMPLE function]
Return true if t is a valid minimal invariant. This is different from constants, in that the specific value of t may not be known at compile time, but it is known that it doesn’t change (e.g., the address of a function local variable).

bool is_gimple_ip_invariant (tree t) [GIMPLE function]
Return true if t is an interprocedural invariant. This means that t is a valid invariant in all functions (e.g. it can be an address of a global variable but not of a local one).

bool is_gimple_ip_invariant_address (tree t) [GIMPLE function]
Return true if t is an ADDR_EXPR that does not change once the program is running (and which is valid in all functions).
12.5.8 Statement validation

bool is_gimple_assign (gimple g)    [GIMPLE function]
  Return true if the code of g is GIMPLE_ASSIGN.

bool is_gimple_call (gimple g)      [GIMPLE function]
  Return true if the code of g is GIMPLE_CALL.

bool is_gimple_debug (gimple g)     [GIMPLE function]
  Return true if the code of g is GIMPLE_DEBUG.

bool gimple_assign_cast_p (gimple g) [GIMPLE function]
  Return true if g is a GIMPLE_ASSIGN that performs a type cast operation.

bool gimple_debug_bind_p (gimple g) [GIMPLE function]
  Return true if g is a GIMPLE_DEBUG that binds the value of an expression to a variable.

12.6 Manipulating GIMPLE statements

This section documents all the functions available to handle each of the GIMPLE instructions.

12.6.1 Common accessors

The following are common accessors for gimple statements.

enum gimple_code gimple_code (gimple g)    [GIMPLE function]
  Return the code for statement G.

basic_block gimple_bb (gimple g)           [GIMPLE function]
  Return the basic block to which statement G belongs to.

tree gimple_block (gimple g)               [GIMPLE function]
  Return the lexical scope block holding statement G.

tree gimple_expr_type (gimple stmt)       [GIMPLE function]
  Return the type of the main expression computed by STMT. Return void_type_node
  if STMT computes nothing. This will only return something meaningful for GIMPLE_ 
  ASSIGN, GIMPLE_COND and GIMPLE_CALL. For all other tuple codes, it will return
  void_type_node.
	enum tree_code gimple_expr_code (gimple stmt) [GIMPLE function]
  Return the tree code for the expression computed by STMT. This is only meaningful
  for GIMPLE_CALL, GIMPLE_ASSIGN and GIMPLE_COND. If STMT is GIMPLE_CALL, it will
  return CALL_EXPR. For GIMPLE_COND, it returns the code of the comparison predicate.
  For GIMPLE_ASSIGN it returns the code of the operation performed by the RHS of the
  assignment.

void gimple_set_block (gimple g, tree block) [GIMPLE function]
  Set the lexical scope block of G to BLOCK.

location_t gimple_locus (gimple g)        [GIMPLE function]
  Return locus information for statement G.
void \texttt{gimple\_set\_locus} (gimple \texttt{g}, location\_t \texttt{locus}) \hfill \text{[GIMPLE function]} \\
Set locus information for statement \texttt{G}.

bool \texttt{gimple\_locus\_empty\_p} (gimple \texttt{g}) \hfill \text{[GIMPLE function]} \\
Return true if \texttt{G} does not have locus information.

bool \texttt{gimple\_no\_warning\_p} (gimple \texttt{stmt}) \hfill \text{[GIMPLE function]} \\
Return true if no warnings should be emitted for statement \texttt{STMT}.

void \texttt{gimple\_set\_visited} (gimple \texttt{stmt}, bool \texttt{visited\_p}) \hfill \text{[GIMPLE function]} \\
Set the visited status on statement \texttt{STMT} to VISITED\_P.

bool \texttt{gimple\_visited\_p} (gimple \texttt{stmt}) \hfill \text{[GIMPLE function]} \\
Return the visited status on statement \texttt{STMT}.

void \texttt{gimple\_set\_plf} (gimple \texttt{stmt}, enum plf\_mask \texttt{plf}, bool \texttt{val\_p}) \hfill \text{[GIMPLE function]} \\
Set pass local flag \texttt{PLF} on statement \texttt{STMT} to VAL\_P.

unsigned int \texttt{gimple\_plf} (gimple \texttt{stmt}, enum plf\_mask \texttt{plf}) \hfill \text{[GIMPLE function]} \\
Return the value of pass local flag \texttt{PLF} on statement \texttt{STMT}.

bool \texttt{gimple\_has\_ops} (gimple \texttt{g}) \hfill \text{[GIMPLE function]} \\
Return true if statement \texttt{G} has register or memory operands.

bool \texttt{gimple\_has\_mem\_ops} (gimple \texttt{g}) \hfill \text{[GIMPLE function]} \\
Return true if statement \texttt{G} has memory operands.

unsigned \texttt{gimple\_num\_ops} (gimple \texttt{g}) \hfill \text{[GIMPLE function]} \\
Return the number of operands for statement \texttt{G}.

tree \* \texttt{gimple\_ops} (gimple \texttt{g}) \hfill \text{[GIMPLE function]} \\
Return the array of operands for statement \texttt{G}.

tree \texttt{gimple\_op} (gimple \texttt{g}, unsigned \texttt{i}) \hfill \text{[GIMPLE function]} \\
Return operand \texttt{I} for statement \texttt{G}.

tree \* \texttt{gimple\_op\_ptr} (gimple \texttt{g}, unsigned \texttt{i}) \hfill \text{[GIMPLE function]} \\
Return a pointer to operand \texttt{I} for statement \texttt{G}.

void \texttt{gimple\_set\_op} (gimple \texttt{g}, unsigned \texttt{i}, tree \texttt{op}) \hfill \text{[GIMPLE function]} \\
Set operand \texttt{I} of statement \texttt{G} to \texttt{OP}.

bitmap \texttt{gimple\_addresses\_taken} (gimple \texttt{stmt}) \hfill \text{[GIMPLE function]} \\
Return the set of symbols that have had their address taken by \texttt{STMT}.

struct def\_optype\_d \* \texttt{gimple\_def\_ops} (gimple \texttt{g}) \hfill \text{[GIMPLE function]} \\
Return the set of DEF operands for statement \texttt{G}.

void \texttt{gimple\_set\_def\_ops} (gimple \texttt{g}, struct def\_optype\_d \* \texttt{def}) \hfill \text{[GIMPLE function]} \\
Set \texttt{DEF} to be the set of DEF operands for statement \texttt{G}.
struct use_optype_d * gimple_use_ops (gimple g)  
Return the set of USE operands for statement G.

void gimple_set_use_ops (gimple g, struct use_optype_d *use)  
Set USE to be the set of USE operands for statement G.

struct voptype_d * gimple_vuse_ops (gimple g)  
Return the set of VUSE operands for statement G.

void gimple_set_vuse_ops (gimple g, struct voptype_d *ops)  
Set OPS to be the set of VUSE operands for statement G.

struct voptype_d * gimple_vdef_ops (gimple g)  
Return the set of VDEF operands for statement G.

void gimple_set_vdef_ops (gimple g, struct voptype_d *ops)  
Set OPS to be the set of VDEF operands for statement G.

bitmap gimple_loaded_syms (gimple g)  
Return the set of symbols loaded by statement G. Each element of the set is the DECL_UID of the corresponding symbol.

bitmap gimple_stored_syms (gimple g)  
Return the set of symbols stored by statement G. Each element of the set is the DECL_UID of the corresponding symbol.

bool gimple_modified_p (gimple g)  
Return true if statement G has operands and the modified field has been set.

bool gimple_has_volatile_ops (gimple stmt)  
Return true if statement STMT contains volatile operands.

void gimple_set_has_volatile_ops (gimple stmt, bool volatilep)  
Return true if statement STMT contains volatile operands.

void update_stmt (gimple s)  
Mark statement S as modified, and update it.

void update_stmt_if_modified (gimple s)  
Update statement S if it has been marked modified.

gimple gimple_copy (gimple stmt)  
Return a deep copy of statement STMT.
12.7 Tuple specific accessors

12.7.1 GIMPLE_ASM

- **gimple gimple_build_asm (const char *string, ninputs, noutputs, ncllobbers, ...)**
  - Build a GIMPLE_ASM statement. This statement is used for building in-line assembly constructs. **STRING** is the assembly code. **NINPUT** is the number of register inputs. **NOUTPUT** is the number of register outputs. **NCLOBBERS** is the number of clobbered registers. The rest of the arguments trees for each input, output, and clobbered registers.

- **gimple gimple_build_asm_vec (const char *, VEC(tree,gc) *, VEC(tree,gc) *, VEC(tree,gc) **) 
  - Identical to gimple_build_asm, but the arguments are passed in VECs.

- **unsigned gimple_asm_ninputs (gimple g)**
  - Return the number of input operands for GIMPLE_ASM G.

- **unsigned gimple_asm_noutputs (gimple g)**
  - Return the number of output operands for GIMPLE_ASM G.

- **unsigned gimple_asm_nclobbers (gimple g)**
  - Return the number of clobber operands for GIMPLE_ASM G.

- **tree gimple_asm_input_op (gimple g, unsigned index)**
  - Return input operand INDEX of GIMPLE_ASM G.

- **void gimple_asm_set_input_op (gimple g, unsigned index, tree in_op)**
  - Set IN_OP to be input operand INDEX in GIMPLE_ASM G.

- **tree gimple_asm_output_op (gimple g, unsigned index)**
  - Return output operand INDEX of GIMPLE_ASM G.

- **void gimple_asm_set_output_op (gimple g, unsigned index, tree out_op)**
  - Set OUT_OP to be output operand INDEX in GIMPLE_ASM G.

- **tree gimple_asm_clobber_op (gimple g, unsigned index)**
  - Return clobber operand INDEX of GIMPLE_ASM G.

- **void gimple_asm_set_clobber_op (gimple g, unsigned index, tree clobber_op)**
  - Set CLOBBER_OP to be clobber operand INDEX in GIMPLE_ASM G.

- **const char * gimple_asm_string (gimple g)**
  - Return the string representing the assembly instruction in GIMPLE_ASM G.

- **bool gimple_asm_volatile_p (gimple g)**
  - Return true if G is an asm statement marked volatile.
void gimple_asm_set_volatile (gimple g)  
Mark asm statement G as volatile.

void gimple_asm_clear_volatile (gimple g)  
Remove volatile marker from asm statement G.

12.7.2 GIMPLE_ASSIGN

gimple gimple_build_assign (tree lhs, tree rhs)  
Build a GIMPLE_ASSIGN statement. The left-hand side is an lvalue passed in lhs. The 
right-hand side can be either a unary or binary tree expression. The expression tree 
rhs will be flattened and its operands assigned to the corresponding operand slots in 
the new statement. This function is useful when you already have a tree expression 
that you want to convert into a tuple. However, try to avoid building expression 
trees for the sole purpose of calling this function. If you already have the operands 
in separate trees, it is better to use gimple_build_assign_with_ops.

gimple gimplify_assign (tree dst, tree src, gimple seq
*pseq_p)  
Build a new GIMPLE_ASSIGN tuple and append it to the end of *SEQ_P.

DST/SRC are the destination and source respectively. You can pass unsimplified trees in 
DST or SRC, in which case they will be converted to a gimple operand if necessary.

This function returns the newly created GIMPLE_ASSIGN tuple.

gimple gimple_build_assign_with_ops (enum tree code 
subcode, tree lhs, tree op1, tree op2)  
This function is similar to gimple_build_assign, but is used to build a GIMPLE_ 
ASSIGN statement when the operands of the right-hand side of the assignment are 
already split into different operands.

The left-hand side is an lvalue passed in lhs. Subcode is the tree_code for the right-
hand side of the assignment. Op1 and op2 are the operands. If op2 is null, subcode 
must be a tree_code for a unary expression.

enum tree_code gimple_assign_rhs_code (gimple g)  
Return the code of the expression computed on the RHS of assignment statement G.

enum gimple_rhs_class gimple_assign_rhs_class 
(gimple g)  
Return the gimple rhs class of the code for the expression computed on the rhs of 
assignment statement G. This will never return GIMPLE_INVALID_RHS.

tree gimple_assign_lhs (gimple g)  
Return the LHS of assignment statement G.

tree * gimple_assign_lhs_ptr (gimple g)  
Return a pointer to the LHS of assignment statement G.

tree gimple_assign_rhs1 (gimple g)  
Return the first operand on the RHS of assignment statement G.
tree * gimple_assign_rhs1_ptr (gimple g)  
Return the address of the first operand on the RHS of assignment statement G.

tree gimple_assign_rhs2 (gimple g)  
Return the second operand on the RHS of assignment statement G.

tree * gimple_assign_rhs2_ptr (gimple g)  
Return the address of the second operand on the RHS of assignment statement G.

tree gimple_assign_rhs3 (gimple g)  
Return the third operand on the RHS of assignment statement G.

tree * gimple_assign_rhs3_ptr (gimple g)  
Return the address of the third operand on the RHS of assignment statement G.

void gimple_assign_set_lhs (gimple g, tree lhs)  
Set LHS to be the LHS operand of assignment statement G.

void gimple_assign_set_rhs1 (gimple g, tree rhs)  
Set RHS to be the first operand on the RHS of assignment statement G.

void gimple_assign_set_rhs2 (gimple g, tree rhs)  
Set RHS to be the second operand on the RHS of assignment statement G.

void gimple_assign_set_rhs3 (gimple g, tree rhs)  
Set RHS to be the third operand on the RHS of assignment statement G.

bool gimple_assign_cast_p (gimple s)  
Return true if S is a type-cast assignment.

12.7.3 GIMPLE_BIND

gimple gimple_build_bind (tree vars, gimple_seq body)  
Build a GIMPLE_BIND statement with a list of variables in VARS and a body of statements in sequence BODY.

tree gimple_bind_vars (gimple g)  
Return the variables declared in the GIMPLE_BIND statement G.

void gimple_bind_set_vars (gimple g, tree vars)  
Set VARS to be the set of variables declared in the GIMPLE_BIND statement G.

void gimple_bind_append_vars (gimple g, tree vars)  
Append VARS to the set of variables declared in the GIMPLE_BIND statement G.

gimple_seq gimple_bind_body (gimple g)  
Return the GIMPLE sequence contained in the GIMPLE_BIND statement G.

void gimple_bind_set_body (gimple g, gimple_seq seq)  
Set SEQ to be sequence contained in the GIMPLE_BIND statement G.

void gimple_bind_add_stmt (gimple gs, gimple stmt)  
Append a statement to the end of a GIMPLE_BIND’s body.
void gimple_bind_add_seq (gimple gs, gimple_seq seq)  
Append a sequence of statements to the end of a GIMPLE_BIND's body.

tree gimple_bind_block (gimple g)  
Return the TREE_BLOCK node associated with GIMPLE_BIND statement G. This is analogous to the BIND_EXPR_BLOCK field in trees.

void gimple_bind_set_block (gimple g, tree block)  
Set BLOCK to be the TREE_BLOCK node associated with GIMPLE_BIND statement G.

12.7.4 GIMPLE_CALL

gimple gimple_build_call (tree fn, unsigned nargs, ...)  
Build a GIMPLE_CALL statement to function FN. The argument FN must be either a FUNCTION_DECL or a gimple call address as determined by is_gimple_call_addr. NARGS are the number of arguments. The rest of the arguments follow the argument NARGS, and must be trees that are valid as rvalues in gimple (i.e., each operand is validated with is_gimple_operand).

gimple gimple_build_call_from_tree (tree call_expr)  
Build a GIMPLE_CALL from a CALL_EXPR node. The arguments and the function are taken from the expression directly. This routine assumes that call_expr is already in GIMPLE form. That is, its operands are GIMPLE values and the function call needs no further simplification. All the call flags in call_expr are copied over to the new GIMPLE_CALL.

void gimple_build_call_vec (tree fn, VEC(tree, heap)*args)  
Identical to gimple_build_call but the arguments are stored in a VEC().

tree gimple_call_lhs (gimple g)  
Return the LHS of call statement G.

void gimple_call_set_lhs (gimple g, tree lhs)  
Set LHS to be the LHS operand of call statement G.

tree gimple_call_fn (gimple g)  
Return the tree node representing the function called by call statement G.

void gimple_call_set_fn (gimple g, tree fn)  
Set FN to be the function called by call statement G. This has to be a gimple value specifying the address of the called function.

void gimple_call_set_fndecl (gimple g, tree fndecl)  
Set the called function to FNDECL.

If a given GIMPLE_CALL's callee is a FUNCTION_DECL, return it. Otherwise return NULL. This function is analogous to get_callee_fndecl in GENERIC.
tree gimple_call_return_type (gimple g)  
Return the type returned by call statement G.

GIMPLE function

tree gimple_call_chain (gimple g)  
Return the static chain for call statement G.

GIMPLE function

void gimple_call_set_chain (gimple g, tree chain)  
Set CHAIN to be the static chain for call statement G.

GIMPLE function

unsigned gimple_call_num_args (gimple g)  
Return the number of arguments used by call statement G.

GIMPLE function

tree gimple_call_arg (gimple g, unsigned index)  
Return the argument at position INDEX for call statement G. The first argument is 0.

GIMPLE function

tree * gimple_call_arg_ptr (gimple g, unsigned index)  
Return a pointer to the argument at position INDEX for call statement G.

GIMPLE function

void gimple_call_set_arg (gimple g, unsigned index, tree arg)  
Set ARG to be the argument at position INDEX for call statement G.

GIMPLE function

void gimple_call_set_tail (gimple s)  
Mark call statement S as being a tail call (i.e., a call just before the exit of a function). These calls are candidate for tail call optimization.

GIMPLE function

bool gimple_call_tail_p (gimple s)  
Return true if GIMPLE_CALL S is marked as a tail call.

GIMPLE function

void gimple_call_mark_uninlinable (gimple s)  
Mark GIMPLE_CALL S as being uninlinable.

GIMPLE function

bool gimple_call_cannot_inline_p (gimple s)  
Return true if GIMPLE_CALL S cannot be inlined.

GIMPLE function

bool gimple_call_noreturn_p (gimple s)  
Return true if S is a noreturn call.

GIMPLE function

gimple gimple_call_copy_skip_args (gimple stmt, bitmap args_to_skip)  
Build a GIMPLE_CALL identical to STMT but skipping the arguments in the positions marked by the set ARGS_TO_SKIP.

GIMPLE function

12.7.5 GIMPLE_CATCH

gimple gimple_build_catch (tree types, gimple_seq handler)  
Build a GIMPLE_CATCH statement. TYPES are the tree types this catch handles. HANDLER is a sequence of statements with the code for the handler.

GIMPLE function

tree gimple_catch_types (gimple g)  
Return the types handled by GIMPLE_CATCH statement G.

GIMPLE function
tree * gimple_catch_types_ptr (gimple g)  
  [GIMPLE function]
  Return a pointer to the types handled by GIMPLE_CATCH statement G.

gimple_seq gimple_catch_handler (gimple g)  
  [GIMPLE function]
  Return the GIMPLE sequence representing the body of the handler of GIMPLE_CATCH statement G.

void gimple_catch_set_types (gimple g, tree t)  
  [GIMPLE function]
  Set T to be the set of types handled by GIMPLE_CATCH G.

void gimple_catch_set_handler (gimple g, gimple_seq handler)  
  [GIMPLE function]
  Set HANDLER to be the body of GIMPLE_CATCH G.

12.7.6 GIMPLE_COND

gimple gimple_build_cond (enum tree_code pred_code, tree lhs, tree rhs, tree t_label, tree f_label)  
  [GIMPLE function]
  Build a GIMPLE_COND statement. A GIMPLE_COND statement compares LHS and RHS and if the condition in PRED_CODE is true, jump to the label in t_label, otherwise jump to the label in f_label. PRED_CODE are relational operator tree codes like EQ_EXPR, LT_EXPR, LE_EXPR, NE_EXPR, etc.

gimple gimple_build_cond_from_tree (tree cond, tree t_label, tree f_label)  
  [GIMPLE function]
  Build a GIMPLE_COND statement from the conditional expression tree COND. T_LABEL and F_LABEL are as in gimple_build_cond.

enum tree_code gimple_cond_code (gimple g)  
  [GIMPLE function]
  Return the code of the predicate computed by conditional statement G.

void gimple_cond_set_code (gimple g, enum tree_code code)  
  [GIMPLE function]
  Set CODE to be the predicate code for the conditional statement G.

tree gimple_cond_lhs (gimple g)  
  [GIMPLE function]
  Return the LHS of the predicate computed by conditional statement G.

void gimple_cond_set_lhs (gimple g, tree lhs)  
  [GIMPLE function]
  Set LHS to be the LHS operand of the predicate computed by conditional statement G.

tree gimple_cond_rhs (gimple g)  
  [GIMPLE function]
  Return the RHS operand of the predicate computed by conditional G.

void gimple_cond_set_rhs (gimple g, tree rhs)  
  [GIMPLE function]
  Set RHS to be the RHS operand of the predicate computed by conditional statement G.

tree gimples cond_true_label (gimple g)  
  [GIMPLE function]
  Return the label used by conditional statement G when its predicate evaluates to true.

void gimple_cond_set_true_label (gimple g, tree label)  
  [GIMPLE function]
  Set LABEL to be the label used by conditional statement G when its predicate evaluates to true.
void gimple_cond_set_false_label (gimple g, tree label)  [GIMPLE function]
Set LABEL to be the label used by conditional statement G when its predicate evaluates
to false.

tree gimple_cond_false_label (gimple g)  [GIMPLE function]
Return the label used by conditional statement G when its predicate evaluates to false.

void gimple_cond_make_false (gimple g)  [GIMPLE function]
Set the conditional COND_STMT to be of the form 'if (1 == 0)'.

void gimple_cond_make_true (gimple g)  [GIMPLE function]
Set the conditional COND_STMT to be of the form 'if (1 == 1)'.

12.7.7 GIMPLE_DEBUG

gimple gimple_build_debug_bind (tree var, tree value,  [GIMPLE function]
gimple stmt)
Build a GIMPLE_DEBUG statement with GIMPLE_DEBUG_BIND of subcode. The effect
of this statement is to tell debug information generation machinery that the value of
user variable var is given by value at that point, and to remain with that value until
var runs out of scope, a dynamically-subsequent debug bind statement overrides the
binding, or conflicting values reach a control flow merge point. Even if components of
the value expression change afterwards, the variable is supposed to retain the same
value, though not necessarily the same location.

It is expected that var be most often a tree for automatic user variables (VAR_DECL
or PARM_DECL) that satisfy the requirements for gimple registers, but it may also be
a tree for a scalarized component of a user variable (ARRAY_REF, COMPONENT_REF), or
a debug temporary (DEBUG_EXPR_DECL).

As for value, it can be an arbitrary tree expression, but it is recommended that it be
in a suitable form for a gimple assignment RHS. It is not expected that user variables
that could appear as var ever appear in value, because in the latter we’d have their
SSA_NAMES instead, but even if they were not in SSA form, user variables appearing
in value are to be regarded as part of the executable code space, whereas those in
var are to be regarded as part of the source code space. There is no way to refer to
the value bound to a user variable within a value expression.

If value is GIMPLE_DEBUG_BIND_NOVALUE, debug information generation machinery is
informed that the variable var is unbound, i.e., that its value is indeterminate, which
sometimes means it is really unavailable, and other times that the compiler could not
keep track of it.

Block and location information for the newly-created stmt are taken from stmt, if
given.

tree gimple_debug_bind_get_var (gimple stmt)  [GIMPLE function]
Return the user variable var that is bound at stmt.

tree gimple_debug_bind_get_value (gimple stmt)  [GIMPLE function]
Return the value expression that is bound to a user variable at stmt.
tree * gimple_debug_bind_get_value_ptr (gimple stmt)  
Return a pointer to the value expression that is bound to a user variable at stmt.

void gimple_debug_bind_set_var (gimple stmt, tree var)  
Modify the user variable bound at stmt to var.

void gimple_debug_bind_set_value (gimple stmt, tree var)  
Modify the value bound to the user variable bound at stmt to value.

void gimple_debug_bind_reset_value (gimple stmt)  
Modify the value bound to the user variable bound at stmt so that the variable becomes unbound.

bool gimple_debug_bind_has_value_p (gimple stmt)  
Return TRUE if stmt binds a user variable to a value, and FALSE if it unbinds the variable.

12.7.8 GIMPLE_EH_FILTER

gimple gimple_build_eh_filter (tree types, gimple_seq failure)  
Build a GIMPLE_EH_FILTER statement. TYPES are the filter’s types. FAILURE is a sequence with the filter’s failure action.

tree gimple_eh_filter_types (gimple g)  
Return the types handled by GIMPLE_EH_FILTER statement G.

tree * gimple_eh_filter_types_ptr (gimple g)  
Return a pointer to the types handled by GIMPLE_EH_FILTER statement G.

gimple_seq gimple_eh_filter_failure (gimple g)  
Return the sequence of statement to execute when GIMPLE_EH_FILTER statement fails.

void gimple_eh_filter_set_types (gimple g, tree types)  
Set TYPES to be the set of types handled by GIMPLE_EH_FILTER G.

void gimple_eh_filter_set_failure (gimple g, gimple_seq failure)  
Set FAILURE to be the sequence of statements to execute on failure for GIMPLE_EH_FILTER G.

bool gimple_eh_filter_must_not_throw (gimple g)  
Return the EH_FILTER_MUST_NOT_THROW flag.

void gimple_eh_filter_set_must_not_throw (gimple g, bool mntp)  
Set the EH_FILTER_MUST_NOT_THROW flag.
12.7.9 GIMPLE_LABEL

gimple gimple_build_label (tree label)  [GIMPLE function]
Build a GIMPLE_LABEL statement with corresponding to the tree label, LABEL.

tree gimple_label_label (gimple g)  [GIMPLE function]
Return the LABEL_DECL node used by GIMPLE_LABEL statement G.

void gimple_label_set_label (gimple g, tree label)  [GIMPLE function]
Set LABEL to be the LABEL_DECL node used by GIMPLE_LABEL statement G.

gimple gimple_build_goto (tree dest)  [GIMPLE function]
Build a GIMPLE_GOTO statement to label DEST.

tree gimple_goto_dest (gimple g)  [GIMPLE function]
Return the destination of the unconditional jump G.

void gimple_goto_set_dest (gimple g, tree dest)  [GIMPLE function]
Set DEST to be the destination of the unconditional jump G.

12.7.10 GIMPLE_NOP

gimple gimple_build_nop (void)  [GIMPLE function]
Build a GIMPLE_NOP statement.

bool gimple_nop_p (gimple g)  [GIMPLE function]
Returns TRUE if statement G is a GIMPLE_NOP.

12.7.11 GIMPLE_OMP_ATOMIC_LOAD

gimple gimple_build_omp_atomic_load (tree lhs, tree rhs)  [GIMPLE function]
Build a GIMPLE_OMP_ATOMIC_LOAD statement. LHS is the left-hand side of the assignment. RHS is the right-hand side of the assignment.

void gimple_omp_atomic_load_set_lhs (gimple g, tree lhs)  [GIMPLE function]
Set the LHS of an atomic load.

tree gimple_omp_atomic_load_lhs (gimple g)  [GIMPLE function]
Get the LHS of an atomic load.

void gimple_omp_atomic_load_set_rhs (gimple g, tree rhs)  [GIMPLE function]
Set the RHS of an atomic set.

tree gimple_omp_atomic_load_rhs (gimple g)  [GIMPLE function]
Get the RHS of an atomic set.

12.7.12 GIMPLE_OMP_ATOMIC_STORE

gimple gimple_build_omp_atomic_store (tree val)  [GIMPLE function]
Build a GIMPLE_OMP_ATOMIC_STORE statement. VAL is the value to be stored.
242 GNU Compiler Collection (GCC) Internals

void gimple_omp_atomic_store_set_val (gimple g, tree val)  
    [GIMPLE function]
    Set the value being stored in an atomic store.

tree gimple_omp_atomic_store_val (gimple g)  
    [GIMPLE function]
    Return the value being stored in an atomic store.

12.7.13 GIMPLE_OMP_CONTINUE

gimple gimple_build_omp_continue (tree control_def, tree control_use)  
    [GIMPLE function]
    Build a GIMPLE_OMP_CONTINUE statement. CONTROL_DEF is the definition of the control variable. CONTROL_USE is the use of the control variable.

tree gimple_omp_continue_control_def (gimple s)  
    [GIMPLE function]
    Return the definition of the control variable on a GIMPLE_OMP_CONTINUE in S.

tree gimple_omp_continue_control_def_ptr (gimple s)  
    [GIMPLE function]
    Same as above, but return the pointer.

tree gimple_omp_continue_set_control_def (gimple s)  
    [GIMPLE function]
    Set the control variable definition for a GIMPLE_OMP_CONTINUE statement in S.

tree gimple_omp_continue_control_use (gimple s)  
    [GIMPLE function]
    Return the use of the control variable on a GIMPLE_OMP_CONTINUE in S.

tree gimple_omp_continue_control_use_ptr (gimple s)  
    [GIMPLE function]
    Same as above, but return the pointer.

tree gimple_omp_continue_set_control_use (gimple s)  
    [GIMPLE function]
    Set the control variable use for a GIMPLE_OMP_CONTINUE statement in S.

12.7.14 GIMPLE_OMP_CRITICAL

gimple gimple_build_omp_critical (gimple_seq body, tree name)  
    [GIMPLE function]
    Build a GIMPLE_OMP_CRITICAL statement. BODY is the sequence of statements for which only one thread can execute. NAME is an optional identifier for this critical block.

tree gimple_omp_critical_name (gimple g)  
    [GIMPLE function]
    Return the name associated with OMP_CRITICAL statement G.

tree * gimple_omp_critical_name_ptr (gimple g)  
    [GIMPLE function]
    Return a pointer to the name associated with OMP critical statement G.

void gimple_omp_critical_set_name (gimple g, tree name)  
    [GIMPLE function]
    Set NAME to be the name associated with OMP critical statement G.
12.7.15 GIMPLE_OMP_FOR

gimple gimple_build_omp_for (gimple_seq body, tree clauses,   [GIMPLE function]
    tree index, tree initial, tree final, tree incr, gimple_seq pre_body, enum
    tree_code omp_for_cond)

Build a GIMPLE_OMP_FOR statement. BODY is sequence of statements inside the for
loop. CLAUSES, are any of the OMP loop construct’s clauses: private, firstprivate,
lastprivate, reductions, ordered, schedule, and nowait. PRE_BODY is the sequence of
statements that are loop invariant. INDEX is the index variable. INITIAL is the initial
value of INDEX. FINAL is final value of INDEX. OMP_FOR_COND is the predicate
used to compare INDEX and FINAL. INCR is the increment expression.

tree gimple_omp_for_clauses (gimple g)   [GIMPLE function]
    Return the clauses associated with OMP_FOR G.

tree * gimple_omp_for_clauses_ptr (gimple g)   [GIMPLE function]
    Return a pointer to the OMP_FOR G.

void gimple_omp_for_set_clauses (gimple g, tree clauses)   [GIMPLE function]
    Set CLAUSES to be the list of clauses associated with OMP_FOR G.

tree gimple_omp_for_index (gimple g)   [GIMPLE function]
    Return the index variable for OMP_FOR G.

tree * gimple_omp_for_index_ptr (gimple g)   [GIMPLE function]
    Return a pointer to the index variable for OMP_FOR G.

void gimple_omp_for_set_index (gimple g, tree index)   [GIMPLE function]
    Set INDEX to be the index variable for OMP_FOR G.

tree gimple_omp_for_initial (gimple g)   [GIMPLE function]
    Return the initial value for OMP_FOR G.

tree * gimple_omp_for_initial_ptr (gimple g)   [GIMPLE function]
    Return a pointer to the initial value for OMP_FOR G.

void gimple_omp_for_set_initial (gimple g, tree initial)   [GIMPLE function]
    Set INITIAL to be the initial value for OMP_FOR G.

tree gimple_omp_for_final (gimple g)   [GIMPLE function]
    Return the final value for OMP_FOR G.

tree * gimple_omp_for_final_ptr (gimple g)   [GIMPLE function]
    Return a pointer to the final value for OMP_FOR G.

void gimple_omp_for_set_final (gimple g, tree final)   [GIMPLE function]
    Set FINAL to be the final value for OMP_FOR G.

tree gimple_omp_for_incr (gimple g)   [GIMPLE function]
    Return the increment value for OMP_FOR G.
tree * gimple_omp_for_incr_ptr (gimple g)  
    [GIMPLE function]
    Return a pointer to the increment value for OMP_FOR G.

void gimple_omp_for_set_incr (gimple g, tree incr)  
    [GIMPLE function]
    Set INCR to be the increment value for OMP_FOR G.

gimple_seq gimple_omp_for_pre_body (gimple g)  
    [GIMPLE function]
    Return the sequence of statements to execute before the OMP_FOR statement G starts.

void gimple_omp_for_set_pre_body (gimple g, gimple_seq pre_body)  
    [GIMPLE function]
    Set PRE_BODY to be the sequence of statements to execute before the OMP_FOR statement G starts.

void gimple_omp_for_set_cond (gimple g, enum tree_code cond)  
    [GIMPLE function]
    Set COND to be the condition code for OMP_FOR G.

elem tree_code gimple_omp_for_cond (gimple g)  
    [GIMPLE function]
    Return the condition code associated with OMP_FOR G.

12.7.16 GIMPLE_OMP_MASTER

gimple gimple_build_omp_master (gimple_seq body)  
    [GIMPLE function]
    Build a GIMPLE_OMP_MASTER statement. BODY is the sequence of statements to be executed by just the master.

12.7.17 GIMPLE_OMP_ORDERED

gimple gimple_build_omp_ordered (gimple_seq body)  
    [GIMPLE function]
    Build a GIMPLE_OMP_ORDERED statement.

    BODY is the sequence of statements inside a loop that will executed in sequence.

12.7.18 GIMPLE_OMP_PARALLEL

gimple gimple_build_omp_parallel (gimple_seq body, tree clauses, tree child_fn, tree data_arg)  
    [GIMPLE function]
    Build a GIMPLE_OMP_PARALLEL statement.

    BODY is sequence of statements which are executed in parallel. CLAUSES, are the OMP parallel construct’s clauses. CHILD_FN is the function created for the parallel threads to execute. DATA_ARG are the shared data argument(s).

bool gimple_omp_parallel_combined_p (gimple g)  
    [GIMPLE function]
    Return true if OMP parallel statement G has the GF_OMP_PARALLEL_COMBINED flag set.

void gimple_omp_parallel_set_combined_p (gimple g)  
    [GIMPLE function]
    Set the GF_OMP_PARALLEL_COMBINED field in OMP parallel statement G.

gimple_seq gimple_omp_body (gimple g)  
    [GIMPLE function]
    Return the body for the OMP statement G.
**void gimple_omp_set_body (gimple g, gimple_seq body)**  
Set BODY to be the body for the OMP statement G.

**tree gimple_omp_parallel_clauses (gimple g)**  
Return the clauses associated with OMP_PARALLEL G.

**tree * gimple_omp_parallel_clauses_ptr (gimple g)**  
Return a pointer to the clauses associated with OMP_PARALLEL G.

**void gimple_omp_parallel_set_clauses (gimple g, tree clauses)**  
Set CLAUSES to be the list of clauses associated with OMP_PARALLEL G.

**tree gimple_omp_parallel_child_fn (gimple g)**  
Return the child function used to hold the body of OMP_PARALLEL G.

**tree * gimple_omp_parallel_child_fn_ptr (gimple g)**  
Return a pointer to the child function used to hold the body of OMP_PARALLEL G.

**void gimple_omp_parallel_set_child_fn (gimple g, tree child_fn)**  
Set CHILD_FN to be the child function for OMP_PARALLEL G.

**tree gimple_omp_parallel_data_arg (gimple g)**  
Return the artificial argument used to send variables and values from the parent to the children threads in OMP_PARALLEL G.

**tree * gimple_omp_parallel_data_arg_ptr (gimple g)**  
Return a pointer to the data argument for OMP_PARALLEL G.

**void gimple_omp_parallel_set_data_arg (gimple g, tree data_arg)**  
Set DATA_ARG to be the data argument for OMP_PARALLEL G.

**bool is_gimple_omp (gimple stmt)**  
Returns true when the gimple statement STMT is any of the OpenMP types.

**12.7.19 GIMPLE_OMP_RETURN**

**gimple gimple_build_omp_return (bool wait_p)**  
Build a GIMPLE_OMP_RETURN statement. WAIT_P is true if this is a non-waiting return.

**void gimple_omp_return_set_nowait (gimple s)**  
Set the nowait flag on GIMPLE_OMP_RETURN statement S.

**bool gimple_omp_return_nowait_p (gimple g)**  
Return true if OMP return statement G has the GF_OMP_RETURN_NOWAIT flag set.
12.7.20 GIMPLE_OMP_SECTION

gimple gimple_build_omp_section (gimple_seq body)  [GIMPLE function]
Build a GIMPLE_OMP_SECTION statement for a sections statement.

BODY is the sequence of statements in the section.

bool gimple_omp_section_last_p (gimple g)  [GIMPLE function]
Return true if OMP section statement G has the GF_OMP_SECTION_LAST flag set.

void gimple_omp_section_set_last (gimple g)  [GIMPLE function]
Set the GF_OMP_SECTION_LAST flag on G.

12.7.21 GIMPLE_OMP_SECTIONS

gimple gimple_build_omp_sections (gimple_seq body, tree clauses)  [GIMPLE function]
Build a GIMPLE_OMP_SECTIONS statement. BODY is a sequence of section statements.
CLAUSES are any of the OMP sections construct’s clauses: private, firstprivate, lastprivate, reduction, and nowait.

gimple gimple_build_omp_sections_switch (void)  [GIMPLE function]
Build a GIMPLE_OMP_SECTIONS_SWITCH statement.

tree gimple_omp_sections_control (gimple g)  [GIMPLE function]
Return the control variable associated with the GIMPLE_OMP_SECTIONS in G.

tree * gimple_omp_sections_control_ptr (gimple g)  [GIMPLE function]
Return a pointer to the clauses associated with the GIMPLE_OMP_SECTIONS in G.

void gimple_omp_sections_set_control (gimple g, tree control)  [GIMPLE function]
Set CONTROL to be the set of clauses associated with the GIMPLE_OMP_SECTIONS in G.

tree gimple_omp_sections_clauses (gimple g)  [GIMPLE function]
Return the clauses associated with OMP_SECTIONS G.

tree * gimple_omp_sections_clauses_ptr (gimple g)  [GIMPLE function]
Return a pointer to the clauses associated with OMP_SECTIONS G.

void gimple_omp_sections_set_clauses (gimple g, tree clauses)  [GIMPLE function]
Set CLAUSES to be the set of clauses associated with OMP_SECTIONS G.

12.7.22 GIMPLE_OMP_SINGLE

 gaping gimple_build_omp_single (gimple_seq body, tree clauses)  [GIMPLE function]
Build a GIMPLE_OMP_SINGLE statement. BODY is the sequence of statements that will be executed once.
CLAUSES are any of the OMP single construct’s clauses: private, firstprivate, copyprivate, nowait.
tree gimple_omp_single_clauses (gimple g)  
   Return the clauses associated with OMP_SINGLE G.

tree * gimple_omp_single_clauses_ptr (gimple g)  
   Return a pointer to the clauses associated with OMP_SINGLE G.

void gimple_omp_single_set_clauses (gimple g, tree clauses)  
   Set CLAUSES to be the clauses associated with OMP_SINGLE G.

12.7.23 GIMPLE_PHI

gimple make_phi_node (tree var, int len)  
   Build a PHI node with len argument slots for variable var.

unsigned gimple_phi_capacity (gimple g)  
   Return the maximum number of arguments supported by GIMPLE_PHI G.

unsigned gimple_phi_num_args (gimple g)  
   Return the number of arguments in GIMPLE_PHI G. This must always be exactly the number of incoming edges for the basic block holding G.

tree gimple_phi_result (gimple g)  
   Return the SSA name created by GIMPLE_PHI G.

tree * gimple_phi_result_ptr (gimple g)  
   Return a pointer to the SSA name created by GIMPLE_PHI G.

void gimple_phi_set_result (gimple g, tree result)  
   Set RESULT to be the SSA name created by GIMPLE_PHI G.

struct phi_arg_d * gimple_phi_arg (gimple g, index)  
   Return the PHI argument corresponding to incoming edge INDEX for GIMPLE_PHI G.

void gimple_phi_set_arg (gimple g, index, struct phi_arg_d * phiarg)  
   Set PHIARG to be the argument corresponding to incoming edge INDEX for GIMPLE_PHI G.

12.7.24 GIMPLE_RESX

gimple gimple_build_resx (int region)  
   Build a GIMPLE_RESX statement which is a statement. This statement is a placeholder for _Unwind_Resume before we know if a function call or a branch is needed. REGION is the exception region from which control is flowing.

int gimple_resx_region (gimple g)  
   Return the region number for GIMPLE_RESX G.

void gimple_resx_set_region (gimple g, int region)  
   Set REGION to be the region number for GIMPLE_RESX G.
12.7.25 GIMPLE_RETURN

**gimple gimple_build_return (tree retval)**  
Build a GIMPLE_RETURN statement whose return value is retval.

**tree gimple_return_retval (gimple g)**  
Return the return value for GIMPLE_RETURN G.

**void gimple_return_set_retval (gimple g, tree retval)**  
Set RETVAL to be the return value for GIMPLE_RETURN G.

12.7.26 GIMPLE_SWITCH

**gimple gimple_build_switch (unsigned nlabels, tree index, tree default_label, ...)**  
Build a GIMPLE_SWITCH statement. NLABELS are the number of labels excluding the default label. The default label is passed in DEFAULT_LABEL. The rest of the arguments are trees representing the labels. Each label is a tree of code CASE_LABEL_EXPR.

**gimple gimple_build_switch_vec (tree index, tree default_label, VEC(tree,heap) *args)**  
This function is an alternate way of building GIMPLE_SWITCH statements. INDEX and DEFAULT_LABEL are as in gimple_build_switch. ARGS is a vector of CASE_LABEL_EXPR trees that contain the labels.

**unsigned gimple_switch_num_labels (gimple g)**  
Return the number of labels associated with the switch statement G.

**void gimple_switch_set_num_labels (gimple g, unsigned nlabels)**  
Set NLABELS to be the number of labels for the switch statement G.

**tree gimple_switch_index (gimple g)**  
Return the index variable used by the switch statement G.

**void gimple_switch_set_index (gimple g, tree index)**  
Set INDEX to be the index variable for switch statement G.

**tree gimple_switch_label (gimple g, unsigned index)**  
Return the label numbered INDEX. The default label is 0, followed by any labels in a switch statement.

**void gimple_switch_set_label (gimple g, unsigned index, tree label)**  
Set the label number INDEX to LABEL. 0 is always the default label.

**tree gimple_switch_default_label (gimple g)**  
Return the default label for a switch statement.

**void gimple_switch_set_default_label (gimple g, tree label)**  
Set the default label for a switch statement.
12.7.27 GIMPLE_TRY

gimple gimple_build_try (gimple_seq eval, gimple_seq cleanup, unsigned int kind)

Build a GIMPLE_TRY statement. EVAL is a sequence with the expression to evaluate. CLEANUP is a sequence of statements to run at clean-up time. KIND is the enumeration value GIMPLE_TRY_CATCH if this statement denotes a try/catch construct or GIMPLE_TRY_FINALLY if this statement denotes a try/finally construct.

enum gimple_try_flags gimple_try_kind (gimple g)

Return the kind of try block represented by GIMPLETRY G. This is either GIMPLE_TRY_CATCH or GIMPLE_TRY_FINALLY.

bool gimple_try_catch_is_cleanup (gimple g)

Return the GIMPLE_TRY_CATCH_IS_CLEANUP flag.

gimple_seq gimple_try_eval (gimple g)

Return the sequence of statements used as the body for GIMPLE_TRY G.

gimple_seq gimple_try_cleanup (gimple g)

Return the sequence of statements used as the cleanup body for GIMPLE_TRY G.

void gimple_try_set_catch_is_cleanup (gimple g, bool catch_is_cleanup)

Set the GIMPLE_TRY_CATCH_IS_CLEANUP flag.

void gimple_try_set_eval (gimple g, gimple_seq eval)

Set EVAL to be the sequence of statements to use as the body for GIMPLETRY G.

void gimple_try_set_cleanup (gimple g, gimple_seq cleanup)

Set CLEANUP to be the sequence of statements to use as the cleanup body for GIMPLETRY G.

12.7.28 GIMPLE_WITH_CLEANUP_EXPR

gimple gimple_build_wce (gimple_seq cleanup)

Build a GIMPLE_WITH_CLEANUP_EXPR statement. CLEANUP is the clean-up expression.

gimple_seq gimple_wce_cleanup (gimple g)

Return the cleanup sequence for cleanup statement G.

void gimple_wce_set_cleanup (gimple g, gimple_seq cleanup)

Set CLEANUP to be the cleanup sequence for G.

bool gimple_wce_cleanup_eh_only (gimple g)

Return the CLEANUP_EH_ONLY flag for a WCE tuple.

void gimple_wce_set_cleanup_eh_only (gimple g, bool eh_only_p)

Set the CLEANUP_EH_ONLY flag for a WCE tuple.
12.8 GIMPLE sequences

GIMPLE sequences are the tuple equivalent of STATEMENT_LIST’s used in GENERIC. They are used to chain statements together, and when used in conjunction with sequence iterators, provide a framework for iterating through statements.

GIMPLE sequences are of type struct gimple_sequence, but are more commonly passed by reference to functions dealing with sequences. The type for a sequence pointer is gimple_seq which is the same as struct gimple_sequence *. When declaring a local sequence, you can define a local variable of type struct gimple_sequence. When declaring a sequence allocated on the garbage collected heap, use the function gimple_seq_alloc documented below.

There are convenience functions for iterating through sequences in the section entitled Sequence Iterators.

Below is a list of functions to manipulate and query sequences.

void gimple_seq_add_stmt (gimple_seq *seq, gimple g) [GIMPLE function]
Link a gimple statement to the end of the sequence *SEQ if G is not NULL. If *SEQ is NULL, allocate a sequence before linking.

void gimple_seq_add_seq (gimple_seq *dest, gimple_seq src) [GIMPLE function]
Append sequence SRC to the end of sequence *DEST if SRC is not NULL. If *DEST is NULL, allocate a new sequence before appending.

gimple_seq gimple_seq_deep_copy (gimple_seq src) [GIMPLE function]
Perform a deep copy of sequence SRC and return the result.

gimple_seq gimple_seq_reverse (gimple_seq seq) [GIMPLE function]
Reverse the order of the statements in the sequence SEQ. Return SEQ.

gimple gimple_seq_first (gimple_seq s) [GIMPLE function]
Return the first statement in sequence S.

gimple gimple_seq_last (gimple_seq s) [GIMPLE function]
Return the last statement in sequence S.

void gimple_seq_set_last (gimple_seq s, gimple last) [GIMPLE function]
Set the last statement in sequence S to the statement in LAST.

void gimple_seq_set_first (gimple_seq s, gimple first) [GIMPLE function]
Set the first statement in sequence S to the statement in FIRST.

void gimple_seq_init (gimple_seq s) [GIMPLE function]
Initialize sequence S to an empty sequence.

gimple_seq gimple_seq_alloc (void) [GIMPLE function]
Allocate a new sequence in the garbage collected store and return it.

void gimple_seq_copy (gimple_seq dest, gimple_seq src) [GIMPLE function]
Copy the sequence SRC into the sequence DEST.
bool gimple_seq_empty_p (gimple_seq s) [GIMPLE function]
Return true if the sequence S is empty.

\[\text{gimple_seq bb_seq}\ (\text{basic_block} \ bb)\] [GIMPLE function]
Returns the sequence of statements in BB.

\[\text{void set_bb_seq}\ (\text{basic_block} \ bb, \ \text{gimple_seq} \ seq)\] [GIMPLE function]
Sets the sequence of statements in BB to SEQ.

bool gimple_seq_singleton_p (gimple_seq seq) [GIMPLE function]
Determine whether SEQ contains exactly one statement.

### 12.9 Sequence iterators

Sequence iterators are convenience constructs for iterating through statements in a sequence. Given a sequence SEQ, here is a typical use of gimple sequence iterators:

```c
  gimple_stmt_iterator gsi;

  for (gsi = gsi_start (seq); !gsi_end_p (gsi); gsi_next (&gsi))
  {
    gimple g = gsi_stmt (gsi);
    /* Do something with gimple statement G. */
  }
```

Backward iterations are possible:
```
  for (gsi = gsi_last (seq); !gsi_end_p (gsi); gsi_prev (&gsi))
```
Forward and backward iterations on basic blocks are possible with gsi_start_bb and gsi_last_bb.

In the documentation below we sometimes refer to enum gsi_iterator_update. The valid options for this enumeration are:

- **GSI_NEW_STMT** Only valid when a single statement is added. Move the iterator to it.
- **GSI_SAME_STMT** Leave the iterator at the same statement.
- **GSI_CONTINUE_LINKING** Move iterator to whatever position is suitable for linking other statements in the same direction.

Below is a list of the functions used to manipulate and use statement iterators.

\[\text{gimple_stmt_iterator gsi_start}\ (\text{gimple_seq} \ seq)\] [GIMPLE function]
Return a new iterator pointing to the sequence SEQ’s first statement. If SEQ is empty, the iterator’s basic block is NULL. Use gsi_start_bb instead when the iterator needs to always have the correct basic block set.

\[\text{gimple_stmt_iterator gsi_start_bb}\ (\text{basic_block} \ bb)\] [GIMPLE function]
Return a new iterator pointing to the first statement in basic block BB.

\[\text{gimple_stmt_iterator gsi_last}\ (\text{gimple_seq} \ seq)\] [GIMPLE function]
Return a new iterator initially pointing to the last statement of sequence SEQ. If SEQ is empty, the iterator’s basic block is NULL. Use gsi_last_bb instead when the iterator needs to always have the correct basic block set.

\[\text{gimple_stmt_iterator gsi_last_bb}\ (\text{basic_block} \ bb)\] [GIMPLE function]
Return a new iterator pointing to the last statement in basic block BB.
bool gsi_end_p (gimple_stmt_iterator i)  
    [GIMPLE function]
    Return TRUE if at the end of i.

bool gsi_one_before_end_p (gimple_stmt_iterator i)  
    [GIMPLE function]
    Return TRUE if we’re one statement before the end of i.

void gsi_next (gimple_stmt_iterator *i)  
    [GIMPLE function]
    Advance the iterator to the next gimple statement.

void gsi_prev (gimple_stmt_iterator *i)  
    [GIMPLE function]
    Advance the iterator to the previous gimple statement.

gimple gsi_stmt (gimple_stmt_iterator i)  
    [GIMPLE function]
    Return the current stmt.

gimple_stmt_iterator gsi_after_labels (basic_block bb)  
    [GIMPLE function]
    Return a block statement iterator that points to the first non-label statement in block BB.

gimple * gsi_stmt_ptr (gimple_stmt_iterator *i)  
    [GIMPLE function]
    Return a pointer to the current stmt.

basic_block gsi_bb (gimple_stmt_iterator i)  
    [GIMPLE function]
    Return the basic block associated with this iterator.

gimple_seq gsi_seq (gimple_stmt_iterator i)  
    [GIMPLE function]
    Return the sequence associated with this iterator.

void gsi_remove (gimple_stmt_iterator *i, bool remove_eh_info)  
    [GIMPLE function]
    Remove the current stmt from the sequence. The iterator is updated to point to the next statement. When REMOVE_EH_INFO is true we remove the statement pointed to by iterator I from the EH tables. Otherwise we do not modify the EH tables. Generally, REMOVE_EH_INFO should be true when the statement is going to be removed from the IL and not reinserted elsewhere.

void gsi_link_seq_before (gimple_stmt_iterator *i,  
                          gimple_seq seq, enum gsi_iterator_update mode)  
    [GIMPLE function]
    Links the sequence of statements SEQ before the statement pointed by iterator I. MODE indicates what to do with the iterator after insertion (see enum gsi_iterator_update above).

void gsi_link_before (gimple_stmt_iterator *i, gimple g,  
                      enum gsi_iterator_update mode)  
    [GIMPLE function]
    Links statement G before the statement pointed-to by iterator I. Updates iterator I according to MODE.

void gsi_link_seq_after (gimple_stmt_iterator *i,  
                         gimple_seq seq, enum gsi_iterator_update mode)  
    [GIMPLE function]
    Links sequence SEQ after the statement pointed-to by iterator I. MODE is as in gsi_insert_after.
void gsi_link_after (gimple_stmt_iterator *i, gimple g, enum gsi_iterator_update mode) [GIMPLE function]
Links statement G after the statement pointed-to by iterator I. MODE is as in gsi_insert_after.

gimple_seq gsi_split_seq_after (gimple_stmt_iterator i) [GIMPLE function]
Move all statements in the sequence after I to a new sequence. Return this new sequence.

gimple_seq gsi_split_seq_before (gimple_stmt_iterator *i) [GIMPLE function]
Move all statements in the sequence before I to a new sequence. Return this new sequence.

void gsi_replace (gimple_stmt_iterator *i, gimple stmt, bool update eh info) [GIMPLE function]
Replace the statement pointed-to by I to STM. If UPDATE_EH_INFO is true, the exception handling information of the original statement is moved to the new statement.

void gsi_insert_before (gimple_stmt_iterator *i, gimple stmt, enum gsi_iterator_update mode) [GIMPLE function]
Insert statement STM before the statement pointed-to by iterator I, update STM’s basic block and scan it for new operands. MODE specifies how to update iterator I after insertion (see enum gsi_iterator_update).

void gsi_insert_seq_before (gimple_stmt_iterator *i, gimple_seq seq, enum gsi_iterator_update mode) [GIMPLE function]
Like gsi_insert_before, but for all the statements in SEQ.

void gsi_insert_after (gimple_stmt_iterator *i, gimple stmt, enum gsi_iterator_update mode) [GIMPLE function]
Insert statement STM after the statement pointed-to by iterator I, update STM’s basic block and scan it for new operands. MODE specifies how to update iterator I after insertion (see enum gsi_iterator_update).

void gsi_insert_seq_after (gimple_stmt_iterator *i, gimple_seq seq, enum gsi_iterator_update mode) [GIMPLE function]
Like gsi_insert_after, but for all the statements in SEQ.

gimple_stmt_iterator gsi_for_stmt (gimple stmt) [GIMPLE function]
Finds iterator for STM.

void gsi_move_after (gimple_stmt_iterator *from, gimple_stmt_iterator *to) [GIMPLE function]
Move the statement at FROM so it comes right after the statement at TO.

void gsi_move_before (gimple_stmt_iterator *from, gimple_stmt_iterator *to) [GIMPLE function]
Move the statement at FROM so it comes right before the statement at TO.

void gsi_move_to_bb_end (gimple_stmt_iterator *from, basic_block bb) [GIMPLE function]
Move the statement at FROM to the end of basic block BB.
void gsi_insert_on_edge (edge e, gimple stmt)  [GIMPLE function]
Add STMT to the pending list of edge E. No actual insertion is made until a call to
gsi_commit_edge_inserts() is made.

void gsi_insert_seq_on_edge (edge e, gimple_seq seq)  [GIMPLE function]
Add the sequence of statements in SEQ to the pending list of edge E. No actual
insertion is made until a call to gsi_commit_edge_inserts() is made.

basic_block gsi_insert_on_edge_immediate (edge e, gimple stmt)  [GIMPLE function]
Similar to gsi_insert_on_edge+gsi_commit_edge_inserts. If a new block has to
be created, it is returned.

void gsi_commit_one_edge_insert (edge e, basic_block *new_bb)  [GIMPLE function]
Commit insertions pending at edge E. If a new block is created, set NEW_BB to this
block, otherwise set it to NULL.

void gsi_commit_edge_inserts (void)  [GIMPLE function]
This routine will commit all pending edge insertions, creating any new basic blocks
which are necessary.

12.10 Adding a new GIMPLE statement code
The first step in adding a new GIMPLE statement code, is modifying the file gimple.def,
which contains all the GIMPLE codes. Then you must add a corresponding structure, and
an entry in union gimple_statement_d, both of which are located in gimple.h. This in
turn, will require you to add a corresponding GTY tag in gsstruct.def, and code to handle
this tag in gss_for_code which is located in gimple.c.

In order for the garbage collector to know the size of the structure you created in
gimple.h, you need to add a case to handle your new GIMPLE statement in gimple_size
which is located in gimple.c.

You will probably want to create a function to build the new gimple statement in
gimple.c. The function should be called gimple_build_new-tuple-name, and should
return the new tuple of type gimple.

If your new statement requires accessors for any members or operands it may have,
put simple inline accessors in gimple.h and any non-trivial accessors in gimple.c with a
corresponding prototype in gimple.h.

12.11 Statement and operand traversals
There are two functions available for walking statements and sequences: walk_gimple_
stmt and walk_gimple_seq, accordingly, and a third function for walking the operands in
a statement: walk_gimple_op.

tree walk_gimple_stmt (gimple_stmt_iterator *gsi,
walk_stmt_fn callback_stmt, walk_tree_fn callback_op, struct walk_stmt_info
*wi)  [GIMPLE function]
This function is used to walk the current statement in GSI, optionally using traversal
state stored in WI. If WI is NULL, no state is kept during the traversal.
The callback `CALLBACK_STMT` is called. If `CALLBACK_STMT` returns true, it means that the callback function has handled all the operands of the statement and it is not necessary to walk its operands.

If `CALLBACK_STMT` is `NULL` or it returns false, `CALLBACK_OP` is called on each operand of the statement via `walk_gimple_op`. If `walk_gimple_op` returns non-`NULL` for any operand, the remaining operands are not scanned.

The return value is that returned by the last call to `walk_gimple_op`, or `NULL_TREE` if no `CALLBACK_OP` is specified.

```c
tree walk_gimple_op (gimple stmt, walk_tree_fn callback_op, struct walk_stmt_info *wi)
```

Use this function to walk the operands of statement `stmt`. Every operand is walked via `walk_tree` with optional state information in `wi`.

`CALLBACK_OP` is called on each operand of `stmt` via `walk_tree`. Additional parameters to `walk_tree` must be stored in `wi`. For each operand `op`, `walk_tree` is called as:

```c
walk_tree (&op, CALLBACK_OP, WI, PSET)
```

If `CALLBACK_OP` returns non-`NULL` for an operand, the remaining operands are not scanned. The return value is that returned by the last call to `walk_tree`, or `NULL_TREE` if no `CALLBACK_OP` is specified.

```c
tree walk_gimple_seq (gimple_seq seq, walk_stmt_fn callback_stmt, walk_tree_fn callback_op, struct walk_stmt_info *wi)
```

This function walks all the statements in the sequence `seq` calling `walk_gimple_stmt` on each one. `wi` is as in `walk_gimple_stmt`. If `walk_gimple_stmt` returns non-`NULL`, the walk is stopped and the value returned. Otherwise, all the statements are walked and `NULL_TREE` returned.
13 Analysis and Optimization of GIMPLE tuples

GCC uses three main intermediate languages to represent the program during compilation: GENERIC, GIMPLE and RTL. GENERIC is a language-independent representation generated by each front end. It is used to serve as an interface between the parser and optimizer. GENERIC is a common representation that is able to represent programs written in all the languages supported by GCC.

GIMPLE and RTL are used to optimize the program. GIMPLE is used for target and language independent optimizations (e.g., inlining, constant propagation, tail call elimination, redundancy elimination, etc). Much like GENERIC, GIMPLE is a language independent, tree based representation. However, it differs from GENERIC in that the GIMPLE grammar is more restrictive: expressions contain no more than 3 operands (except function calls), it has no control flow structures and expressions with side-effects are only allowed on the right hand side of assignments. See the chapter describing GENERIC and GIMPLE for more details.

This chapter describes the data structures and functions used in the GIMPLE optimizers (also known as “tree optimizers” or “middle end”). In particular, it focuses on all the macros, data structures, functions and programming constructs needed to implement optimization passes for GIMPLE.

13.1 Annotations

The optimizers need to associate attributes with variables during the optimization process. For instance, we need to know whether a variable has aliases. All these attributes are stored in data structures called annotations which are then linked to the field `ann` in `struct tree_common`.

Presently, we define annotations for variables (`var_ann_t`). Annotations are defined and documented in ‘tree-flow.h’.

13.2 SSA Operands

Almost every GIMPLE statement will contain a reference to a variable or memory location. Since statements come in different shapes and sizes, their operands are going to be located at various spots inside the statement’s tree. To facilitate access to the statement’s operands, they are organized into lists associated inside each statement’s annotation. Each element in an operand list is a pointer to a `VAR_DECL`, `PARM_DECL` or `SSA_NAME` tree node. This provides a very convenient way of examining and replacing operands.

Data flow analysis and optimization is done on all tree nodes representing variables. Any node for which `SSA_VAR_P` returns nonzero is considered when scanning statement operands. However, not all `SSA_VAR_P` variables are processed in the same way. For the purposes of optimization, we need to distinguish between references to local scalar variables and references to globals, statics, structures, arrays, aliased variables, etc. The reason is simple, the compiler can gather complete data flow information for a local scalar. On the other hand, a global variable may be modified by a function call, it may not be possible to keep track of all the elements of an array or the fields of a structure, etc.

The operand scanner gathers two kinds of operands: `real` and `virtual`. An operand for which `is_gimple_reg` returns true is considered real, otherwise it is a virtual operand. We
also distinguish between uses and definitions. An operand is used if its value is loaded by
the statement (e.g., the operand at the RHS of an assignment). If the statement assigns a
new value to the operand, the operand is considered a definition (e.g., the operand at the
LHS of an assignment).

Virtual and real operands also have very different data flow properties. Real operands
are unambiguous references to the full object that they represent. For instance, given

```
{  
    int a, b;  
    a = b  
}
```

Since `a` and `b` are non-aliased locals, the statement `a = b` will have one real definition and
one real use because variable `a` is completely modified with the contents of variable `b`. Real
definition are also known as killing definitions. Similarly, the use of `b` reads all its bits.

In contrast, virtual operands are used with variables that can have a partial or ambiguous
reference. This includes structures, arrays, globals, and aliased variables. In these cases, we
have two types of definitions. For globals, structures, and arrays, we can determine from
a statement whether a variable of these types has a killing definition. If the variable does,
then the statement is marked as having a must definition of that variable. However, if a
statement is only defining a part of the variable (i.e. a field in a structure), or if we know
that a statement might define the variable but we cannot say for sure, then we mark that
statement as having a may definition. For instance, given

```
{  
    int a, b, *p;  
    if (...)  
        p = &a;  
    else  
        p = &b;  
    *p = 5;  
    return *p;  
}
```

The assignment `*p = 5` may be a definition of `a` or `b`. If we cannot determine statically
where `p` is pointing to at the time of the store operation, we create virtual definitions to
mark that statement as a potential definition site for `a` and `b`. Memory loads are similarly
marked with virtual use operands. Virtual operands are shown in tree dumps right before
the statement that contains them. To request a tree dump with virtual operands, use the
`-vops` option to `-fdump-tree`:

```
{  
    int a, b, *p;  
    if (...)  
        p = &a;  
    else  
        p = &b;  
    # a = VDEF <a>  
    # b = VDEF <b>  
    *p = 5;  
    # VUSE <a>  
    # VUSE <b>  
    return *p;  
}`
Notice that VDEF operands have two copies of the referenced variable. This indicates that this is not a killing definition of that variable. In this case we refer to it as a *may definition* or *aliased store*. The presence of the second copy of the variable in the VDEF operand will become important when the function is converted into SSA form. This will be used to link all the non-killing definitions to prevent optimizations from making incorrect assumptions about them.

Operands are updated as soon as the statement is finished via a call to `update_stmt`. If statement elements are changed via `SET_USE` or `SET_DEF`, then no further action is required (i.e., those macros take care of updating the statement). If changes are made by manipulating the statement’s tree directly, then a call must be made to `update_stmt` when complete. Calling one of the `bsi_insert` routines or `bsi_replace` performs an implicit call to `update_stmt`.

13.2.1 Operand Iterators And Access Routines

Operands are collected by `tree-ssa-operands.c`. They are stored inside each statement’s annotation and can be accessed through either the operand iterators or an access routine.

The following access routines are available for examining operands:

1. **SINGLE_SSA_{USE,DEF,TREE}_OPERAND**: These accessors will return NULL unless there is exactly one operand matching the specified flags. If there is exactly one operand, the operand is returned as either a `tree`, `def_operand_p`, or `use_operand_p`.
   ```c
   tree t = SINGLE_SSA_TREE_OPERAND (stmt, flags);
   use_operand_p u = SINGLE_SSA_USE_OPERAND (stmt, SSA_ALL_VIRTUAL_USES);
   def_operand_p d = SINGLE_SSA_DEF_OPERAND (stmt, SSA_OP_ALL_DEFS);
   ```

2. **ZERO_SSA_OPERANDS**: This macro returns true if there are no operands matching the specified flags.
   ```c
   if (ZERO_SSA_OPERANDS (stmt, SSA_OP_ALL_VIRTUALS))
     return;
   ```

3. **NUM_SSA_OPERANDS**: This macro Returns the number of operands matching ‘flags’. This actually executes a loop to perform the count, so only use this if it is really needed.
   ```c
   int count = NUM_SSA_OPERANDS (stmt, flags)
   ```

If you wish to iterate over some or all operands, use the **FOR_EACH_SSA_{USE,DEF,TREE}_OPERAND** iterator. For example, to print all the operands for a statement:

```c
void print_ops (tree stmt)
{
  ssa_op_iter;
  tree var;

  FOR_EACH_SSA_TREE_OPERAND (var, stmt, iter, SSA_OP_ALL_OPERANDS)
    print_generic_expr (stderr, var, TDF_SLIM);
}
```

How to choose the appropriate iterator:

1. Determine whether you are need to see the operand pointers, or just the trees, and choose the appropriate macro:

<table>
<thead>
<tr>
<th>Need</th>
<th>Macro:</th>
</tr>
</thead>
<tbody>
<tr>
<td>----</td>
<td>--------</td>
</tr>
</tbody>
</table>


2. You need to declare a variable of the type you are interested in, and an ssa_op_iter structure which serves as the loop controlling variable.

3. Determine which operands you wish to use, and specify the flags of those you are interested in. They are documented in `tree-ssa-operands.h`:

   #define SSA_OP_USE 0x01 /* Real USE operands. */
   #define SSA_OP_DEF 0x02 /* Real DEF operands. */
   #define SSA_OP_VUSE 0x04 /* VUSE operands. */
   #define SSA_OP_VMAYUSE 0x08 /* USE portion of VDEFS. */
   #define SSA_OP_VDEF 0x10 /* DEF portion of VDEFS. */

   /* These are commonly grouped operand flags. */
   #define SSA_OP_VIRTUAL_USES (SSA_OP_VUSE | SSA_OP_VMAYUSE)
   #define SSA_OP_VIRTUAL_DEFS (SSA_OP_VDEF)
   #define SSA_OP_ALL_USES (SSA_OP_VIRTUAL_USES | SSA_OP_USE)
   #define SSA_OP_ALL_DEFS (SSA_OP_VIRTUAL_DEFS | SSA_OP_DEF)
   #define SSA_OP_ALL_OPERANDS (SSA_OP_ALL_USES | SSA_OP_ALL_DEFS)

So if you want to look at the use pointers for all the USE and VUSE operands, you would do something like:

```c
use_operand_p use_p;
ssa_op_iter iter;

FOR_EACH_SSA_USE_OPERAND (use_p, stmt, iter, (SSA_OP_USE | SSA_OP_VUSE))
{
    process_use_ptr (use_p);
}
```

The TREE macro is basically the same as the USE and DEF macros, only with the use or def dereferenced via USE_FROM_PTR (use_p) and DEF_FROM_PTR (def_p). Since we aren’t using operand pointers, use and defs flags can be mixed.

```c
tree var;
ssa_op_iter iter;

FOR_EACH_SSA_TREE_OPERAND (var, stmt, iter, SSA_OP_VUSE)
{
    print_generic_expr (stderr, var, TDF_SLIM);
}
```

VDEFs are broken into two flags, one for the DEF portion (SSA_OP_VDEF) and one for the USE portion (SSA_OP_VMAYUSE). If all you want to look at are the VDEFs together, there is a fourth iterator macro for this, which returns both a def_operand_p and a use_operand_p for each VDEF in the statement. Note that you don’t need any flags for this one.

```c
use_operand_p use_p;
def_operand_p def_p;
ssa_op_iter iter;

FOR_EACH_SSA_MAYDEF_OPERAND (def_p, use_p, stmt, iter)
{
    my_code;
}
```

There are many examples in the code as well, as well as the documentation in `tree-ssa-operands.h`. 
There are also a couple of variants on the stmt iterators regarding PHI nodes.

`FOR_EACH_PHI_ARG` Works exactly like `FOR_EACH_SSA_USE_OPERAND`, except it works over PHI arguments instead of statement operands.

```c
/* Look at every virtual PHI use. */
FOR_EACH_PHI_ARG (use_p, phi_stmt, iter, SSA_OP_VIRTUAL_USES)
{
    my_code;
}

/* Look at every real PHI use. */
FOR_EACH_PHI_ARG (use_p, phi_stmt, iter, SSA_OP_USES)
    my_code;

/* Look at every PHI use. */
FOR_EACH_PHI_ARG (use_p, phi_stmt, iter, SSA_OP_ALL_USES)
    my_code;
```

`FOR_EACH_PHI_OR_STMT_{USE,DEF}` works exactly like `FOR_EACH_SSA_{USE,DEF}_OPERAND`, except it will function on either a statement or a PHI node. These should be used when it is appropriate but they are not quite as efficient as the individual `FOR_EACH_PHI` and `FOR_EACH_SSA` routines.

```c
FOR_EACH_PHI_OR_STMT_{USE,DEF} (use_operand_p, stmt, iter, flags)
{
    my_code;
}
```

### 13.2.2 Immediate Uses

Immediate use information is now always available. Using the immediate use iterators, you may examine every use of any `SSA_NAME`. For instance, to change each use of `ssa_var` to `ssa_var2` and call `fold_stmt` on each stmt after that is done:

```c
use_operand_p imm_use_p;
imm_use_iterator iterator;
tree ssa_var, stmt;

FOR_EACH_IMM_USE_STMT (stmt, iterator, ssa_var)
{
    FOR_EACH_IMM_USE_ON_STMT (imm_use_p, iterator)
        SET_USE (imm_use_p, ssa_var_2);
        fold_stmt (stmt);
}
```

There are 2 iterators which can be used. `FOR_EACH_IMM_USE_FAST` is used when the immediate uses are not changed, i.e., you are looking at the uses, but not setting them.

If they do get changed, then care must be taken that things are not changed under the iterators, so use the `FOR_EACH_IMM_USE_STMT` and `FOR_EACH_IMM_USE_ON_STMT` iterators. They attempt to preserve the sanity of the use list by moving all the uses for a statement into a controlled position, and then iterating over those uses. Then the optimization can manipulate the stmt when all the uses have been processed. This is a little slower than the
FAST version since it adds a placeholder element and must sort through the list a bit for each statement. This placeholder element must be also be removed if the loop is terminated early. The macro `BREAK_FROM_IMM_USE_SAFE` is provided to do this:

```c
FOR_EACH_IMM_USE_STMT (stmt, iterator, ssa_var)
{
    if (stmt == last_stmt)
        BREAK_FROM_SAFE_IMM_USE (iter);

    FOR_EACH_IMM_USE_ON_STMT (imm_use_p, iterator)
        SET_USE (imm_use_p, ssa_var_2);
    fold_stmt (stmt);
}
```

There are checks in `verify_ssa` which verify that the immediate use list is up to date, as well as checking that an optimization didn't break from the loop without using this macro. It is safe to simply 'break'; from a `FOR_EACH_IMM_USE_FAST` traverse.

Some useful functions and macros:

1. `has_zero_uses (ssa_var)` : Returns true if there are no uses of `ssa_var`.
2. `has_single_use (ssa_var)` : Returns true if there is only a single use of `ssa_var`.
3. `single_imm_use (ssa_var, use_operand_p *ptr, tree *stmt)` : Returns true if there is only a single use of `ssa_var`, and also returns the use pointer and statement it occurs in, in the second and third parameters.
4. `num_imm_uses (ssa_var)` : Returns the number of immediate uses of `ssa_var`. It is better not to use this if possible since it simply utilizes a loop to count the uses.
5. `PHI_ARG_INDEX_FROM_USE (use_p)` : Given a use within a PHI node, return the index number for the use. An assert is triggered if the use isn’t located in a PHI node.
6. `USE_STMT (use_p)` : Return the statement a use occurs in.

Note that uses are not put into an immediate use list until their statement is actually inserted into the instruction stream via a `bsi_*` routine.

It is also still possible to utilize lazy updating of statements, but this should be used only when absolutely required. Both alias analysis and the dominator optimizations currently do this.

When lazy updating is being used, the immediate use information is out of date and cannot be used reliably. Lazy updating is achieved by simply marking statements modified via calls to `mark_stmt_modified` instead of `update_stmt`. When lazy updating is no longer required, all the modified statements must have `update_stmt` called in order to bring them up to date. This must be done before the optimization is finished, or `verify_ssa` will trigger an abort.

This is done with a simple loop over the instruction stream:

```c
block_stmt_iterator bsi;
basic_block bb;
FOR_EACH_BB (bb)
{
    for (bsi = bsi_start (bb); !bsi_end_p (bsi); bsi_next (&bsi))
        update_stmt_if_modified (bsi_stmt (bsi));
}
```
13.3 Static Single Assignment

Most of the tree optimizers rely on the data flow information provided by the Static Single Assignment (SSA) form. We implement the SSA form as described in R. Cytron, J. Ferrante, B. Rosen, M. Wegman, and K. Zadeck. Efficiently Computing Static Single Assignment Form and the Control Dependence Graph. ACM Transactions on Programming Languages and Systems, 13(4):451-490, October 1991.

The SSA form is based on the premise that program variables are assigned in exactly one location in the program. Multiple assignments to the same variable create new versions of that variable. Naturally, actual programs are seldom in SSA form initially because variables tend to be assigned multiple times. The compiler modifies the program representation so that every time a variable is assigned in the code, a new version of the variable is created. Different versions of the same variable are distinguished by subscripting the variable name with its version number. Variables used in the right-hand side of expressions are renamed so that their version number matches that of the most recent assignment.

We represent variable versions using SSA_NAME nodes. The renaming process in 'tree-ssa.c' wraps every real and virtual operand with an SSA_NAME node which contains the version number and the statement that created the SSA_NAME. Only definitions and virtual definitions may create new SSA_NAME nodes.

Sometimes, flow of control makes it impossible to determine the most recent version of a variable. In these cases, the compiler inserts an artificial definition for that variable called PHI function or PHI node. This new definition merges all the incoming versions of the variable to create a new name for it. For instance,

```
if (...) 
    a_1 = 5;
else if (...) 
    a_2 = 2;
else 
    a_3 = 13;
# a_4 = PHI <a_1, a_2, a_3>
return a_4;
```

Since it is not possible to determine which of the three branches will be taken at runtime, we don’t know which of a_1, a_2 or a_3 to use at the return statement. So, the SSA renamer creates a new version a_4 which is assigned the result of “merging” a_1, a_2 and a_3. Hence, PHI nodes mean “one of these operands. I don’t know which”.

The following macros can be used to examine PHI nodes

\textbf{PHI\_RESULT (phi)} \hspace{1cm} \textbf{[Macro]}
\begin{itemize}
  \item Returns the SSA_NAME created by PHI node \textit{phi} (i.e., \textit{phi}'s LHS).
\end{itemize}

\textbf{PHI\_NUM\_ARGS (phi)} \hspace{1cm} \textbf{[Macro]}
\begin{itemize}
  \item Returns the number of arguments in \textit{phi}. This number is exactly the number of incoming edges to the basic block holding \textit{phi}.
\end{itemize}

\textbf{PHI\_ARG\_ELT (phi, i)} \hspace{1cm} \textbf{[Macro]}
\begin{itemize}
  \item Returns a tuple representing the \textit{i}th argument of \textit{phi}. Each element of this tuple contains an SSA_NAME \textit{var} and the incoming edge through which \textit{var} flows.
\end{itemize}
PHI_ARG_EDGE (phi, i)  \[\text{Macro}\]

Returns the incoming edge for the $i$th argument of $phi$.

PHI_ARG_DEF (phi, i)  \[\text{Macro}\]

Returns the SSA_NAME for the $i$th argument of $phi$.

### 13.3.1 Preserving the SSA form

Some optimization passes make changes to the function that invalidate the SSA property. This can happen when a pass has added new symbols or changed the program so that variables that were previously aliased aren’t anymore. Whenever something like this happens, the affected symbols must be renamed into SSA form again. Transformations that emit new code or replicate existing statements will also need to update the SSA form.

Since GCC implements two different SSA forms for register and virtual variables, keeping the SSA form up to date depends on whether you are updating register or virtual names. In both cases, the general idea behind incremental SSA updates is similar: when new SSA names are created, they typically are meant to replace other existing names in the program.

For instance, given the following code:

```plaintext
1 L0:
2 x_1 = PHI (0, x_5)
3 if (x_1 < 10)
4   if (x_1 > 7)
5     y_2 = 0
6   else
7     y_3 = x_1 + x_7
8   endif
9   x_5 = x_1 + 1
10  goto L0;
11 endif
```

Suppose that we insert new names $x_{10}$ and $x_{11}$ (lines 4 and 8).

```plaintext
1 L0:
2 x_{10} = PHI (0, x_5)
3 if (x_{10} < 10)
4   x_{10} = ...
5   if (x_{10} > 7)
6     y_2 = 0
7   else
8     x_{11} = ...
9     y_3 = x_{10} + x_7
10   endif
11   x_5 = x_{11} + 1
12  goto L0;
13 endif
```

We want to replace all the uses of $x_1$ with the new definitions of $x_{10}$ and $x_{11}$. Note that the only uses that should be replaced are those at lines 5, 9 and 11. Also, the use of $x_7$ at line 9 should not be replaced (this is why we cannot just mark symbol $x$ for renaming).

Additionally, we may need to insert a PHI node at line 11 because that is a merge point for $x_{10}$ and $x_{11}$. So the use of $x_1$ at line 11 will be replaced with the new PHI node. The insertion of PHI nodes is optional. They are not strictly necessary to preserve the SSA form, and depending on what the caller inserted, they may not even be useful for the optimizers.
Updating the SSA form is a two step process. First, the pass has to identify which names need to be updated and/or which symbols need to be renamed into SSA form for the first time. When new names are introduced to replace existing names in the program, the mapping between the old and the new names are registered by calling `register_new_name_mapping` (note that if your pass creates new code by duplicating basic blocks, the call to `tree_duplicate_bb` will set up the necessary mappings automatically). On the other hand, if your pass exposes a new symbol that should be put in SSA form for the first time, the new symbol should be registered with `mark_sym_for_renaming`.

After the replacement mappings have been registered and new symbols marked for renaming, a call to `update_ssa` makes the registered changes. This can be done with an explicit call or by creating TODO flags in the `tree_opt_pass` structure for your pass. There are several TODO flags that control the behavior of `update_ssa`:

- **TODO_update_ssa**. Update the SSA form inserting PHI nodes for newly exposed symbols and virtual names marked for updating. When updating real names, only insert PHI nodes for a real name $O_j$ in blocks reached by all the new and old definitions for $O_j$. If the iterated dominance frontier for $O_j$ is not pruned, we may end up inserting PHI nodes in blocks that have one or more edges with no incoming definition for $O_j$. This would lead to uninitialized warnings for $O_j$’s symbol.

- **TODO_update_ssa_no_phi**. Update the SSA form without inserting any new PHI nodes at all. This is used by passes that have either inserted all the PHI nodes themselves or passes that need only to patch use-def and def-def chains for virtuals (e.g., DCE).

- **TODO_update_ssa_full_phi**. Insert PHI nodes everywhere they are needed. No pruning of the IDF is done. This is used by passes that need the PHI nodes for $O_j$ even if it means that some arguments will come from the default definition of $O_j$’s symbol (e.g., `pass_linear_transform`).

  WARNING: If you need to use this flag, chances are that your pass may be doing something wrong. Inserting PHI nodes for an old name where not all edges carry a new replacement may lead to silent codegen errors or spurious uninitialized warnings.

- **TODO_update_ssa_only.virtuals**. Passes that update the SSA form on their own may want to delegate the updating of virtual names to the generic updater. Since FUD chains are easier to maintain, this simplifies the work they need to do. NOTE: If this flag is used, any OLD->NEW mappings for real names are explicitly destroyed and only the symbols marked for renaming are processed.

### 13.3.2 Preserving the virtual SSA form

The virtual SSA form is harder to preserve than the non-virtual SSA form mainly because the set of virtual operands for a statement may change at what some would consider unexpected times. In general, statement modifications should be bracketed between calls to `push_stmt_changes` and `pop_stmt_changes`. For example,

```c
munge_stmt (tree stmt)
{
    push_stmt_changes (&stmt);
    ... rewrite STMT ...
    pop_stmt_changes (&stmt);
}
```
The call to `push_stmt_changes` saves the current state of the statement operands and the call to `pop_stmt_changes` compares the saved state with the current one and does the appropriate symbol marking for the SSA renamer.

It is possible to modify several statements at a time, provided that `push_stmt_changes` and `pop_stmt_changes` are called in LIFO order, as when processing a stack of statements.

Additionally, if the pass discovers that it did not need to make changes to the statement after calling `push_stmt_changes`, it can simply discard the topmost change buffer by calling `discard_stmt_changes`. This will avoid the expensive operand re-scan operation and the buffer comparison that determines if symbols need to be marked for renaming.

### 13.3.3 Examining SSA_NAME nodes

The following macros can be used to examine SSA_NAME nodes

```c
SSA_NAME_DEF_STMT (var)
```

Returns the statement \( s \) that creates the SSA_NAME \( var \). If \( s \) is an empty statement (i.e., \( IS_EMPTY_STMT(s) \) returns `true`), it means that the first reference to this variable is a USE or a VUSE.

```c
SSA_NAME_VERSION (var)
```

Returns the version number of the SSA_NAME object \( var \).

### 13.3.4 Walking use-def chains

```c
void walk_use_def_chains (var, fn, data)
```

Walks use-def chains starting at the SSA_NAME node \( var \). Calls function \( fn \) at each reaching definition found. Function \( FN \) takes three arguments: \( var \), its defining statement (\( def_stmt \)) and a generic pointer to whatever state information that \( fn \) may want to maintain (\( data \)). Function \( fn \) is able to stop the walk by returning `true`, otherwise in order to continue the walk, \( fn \) should return `false`.

Note, that if \( def_stmt \) is a PHI node, the semantics are slightly different. For each argument \( arg \) of the PHI node, this function will:

1. Walk the use-def chains for \( arg \).
2. Call \( FN (arg, phi, data) \).

Note how the first argument to \( fn \) is no longer the original variable \( var \), but the PHI argument currently being examined. If \( fn \) wants to get at \( var \), it should call `PHI_RESULT (phi)`.

### 13.3.5 Walking the dominator tree

```c
void walk_dominator_tree (walk_data, bb)
```

This function walks the dominator tree for the current CFG calling a set of callback functions defined in `struct dom_walk_data` in ‘domwalk.h’. The call back functions you need to define give you hooks to execute custom code at various points during traversal:

1. Once to initialize any local data needed while processing \( bb \) and its children.
   This local data is pushed into an internal stack which is automatically pushed and popped as the walker traverses the dominator tree.
2. Once before traversing all the statements in the bb.
3. Once for every statement inside bb.
4. Once after traversing all the statements and before recursing into bb’s dominator children.
5. It then recurses into all the dominator children of bb.
6. After recursing into all the dominator children of bb it can, optionally, traverse every statement in bb again (i.e., repeating steps 2 and 3).
7. Once after walking the statements in bb and bb’s dominator children. At this stage, the block local data stack is popped.

### 13.4 Alias analysis

Alias analysis in GIMPLE SSA form consists of two pieces. First the virtual SSA web ties conflicting memory accesses and provides a SSA use-def chain and SSA immediate-use chains for walking possibly dependent memory accesses. Second an alias-oracle can be queried to disambiguate explicit and implicit memory references.

1. Memory SSA form.

   All statements that may use memory have exactly one accompanied use of a virtual SSA name that represents the state of memory at the given point in the IL. All statements that may define memory have exactly one accompanied definition of a virtual SSA name using the previous state of memory and defining the new state of memory after the given point in the IL.

   ```c
   int i;
   int foo (void)
   {
     # .MEM_3 = VDEF <.MEM_2(D)>
     i = 1;
     # VUSE <.MEM_3>
     return i;
   }
   ```

   The virtual SSA names in this case are `.MEM_2(D)` and `.MEM_3`. The store to the global variable `i` defines `.MEM_3` invalidating `.MEM_2(D)`. The load from `i` uses that new state `.MEM_3`.

   The virtual SSA web serves as constraints to SSA optimizers preventing illegitimate code-motion and optimization. It also provides a way to walk related memory statements.

2. Points-to and escape analysis.

   Points-to analysis builds a set of constraints from the GIMPLE SSA IL representing all pointer operations and facts we do or do not know about pointers. Solving this set of constraints yields a conservatively correct solution for each pointer variable in the program (though we are only interested in SSA name pointers) as to what it may possibly point to.

   This points-to solution for a given SSA name pointer is stored in the `pt_solution` sub-structure of the `SSA_NAME_PTR_INFO` record. The following accessor functions are available:
   - `pt_solution_includes`
- **pt_solutions_intersect**

  Points-to analysis also computes the solution for two special set of pointers, **ESCAPED** and **CALLUSED**. Those represent all memory that has escaped the scope of analysis or that is used by pure or nested const calls.

3. **Type-based alias analysis**

   Type-based alias analysis is frontend dependent though generic support is provided by the middle-end in **alias.c**. TBAA code is used by both tree optimizers and RTL optimizers.

   Every language that wishes to perform language-specific alias analysis should define a function that computes, given a **tree** node, an alias set for the node. Nodes in different alias sets are not allowed to alias. For an example, see the C front-end function **c_get_alias_set**.

4. **Tree alias-oracle**

   The tree alias-oracle provides means to disambiguate two memory references and memory references against statements. The following queries are available:

   - **refs_may_alias_p**
   - **ref_maybe_used_byStmt_p**
   - **stmt_may_clobber_ref_p**

   In addition to those two kind of statement walkers are available walking statements related to a reference ref. **walk_non_alias_vuses** walks over dominating memory defining statements and calls back if the statement does not clobber ref providing the non-aliased VUSE. The walk stops at the first clobbering statement or if asked to. **walk_aliased_vdefs** walks over dominating memory defining statements and calls back on each statement clobbering ref providing its aliasing VDEF. The walk stops if asked to.

13.5 **Memory model**

   The memory model used by the middle-end models that of the C/C++ languages. The middle-end has the notion of an effective type of a memory region which is used for type-based alias analysis.

   The following is a refinement of ISO C99 6.5/6, clarifying the block copy case to follow common sense and extending the concept of a dynamic effective type to objects with a declared type as required for C++.

   The effective type of an object for an access to its stored value is the declared type of the object or the effective type determined by a previous store to it. If a value is stored into an object through an lvalue having a type that is not a character type, then the type of the lvalue becomes the effective type of the object for that access and for subsequent accesses that do not modify the stored value. If a value is copied into an object using memcpy or memmove, or is copied as an array of character type, then the effective type of the modified object for that access and for subsequent accesses that do not modify the value is undetermined. For all other accesses to an object, the effective type of the object is simply the type of the lvalue used for the access.
14 Analysis and Representation of Loops

GCC provides extensive infrastructure for work with natural loops, i.e., strongly connected components of CFG with only one entry block. This chapter describes representation of loops in GCC, both on GIMPLE and in RTL, as well as the interfaces to loop-related analyses (induction variable analysis and number of iterations analysis).

14.1 Loop representation

This chapter describes the representation of loops in GCC, and functions that can be used to build, modify and analyze this representation. Most of the interfaces and data structures are declared in `cfgloop.h`. At the moment, loop structures are analyzed and this information is updated only by the optimization passes that deal with loops, but some efforts are being made to make it available throughout most of the optimization passes.

In general, a natural loop has one entry block (header) and possibly several back edges (latches) leading to the header from the inside of the loop. Loops with several latches may appear if several loops share a single header, or if there is a branching in the middle of the loop. The representation of loops in GCC however allows only loops with a single latch. During loop analysis, headers of such loops are split and forwarder blocks are created in order to disambiguate their structures. Heuristic based on profile information and structure of the induction variables in the loops is used to determine whether the latches correspond to sub-loops or to control flow in a single loop. This means that the analysis sometimes changes the CFG, and if you run it in the middle of an optimization pass, you must be able to deal with the new blocks. You may avoid CFG changes by passing LOOPS_MAY_HAVE_MULTIPLE_LATCHES flag to the loop discovery, note however that most other loop manipulation functions will not work correctly for loops with multiple latch edges (the functions that only query membership of blocks to loops and subloop relationships, or enumerate and test loop exits, can be expected to work).

Body of the loop is the set of blocks that are dominated by its header, and reachable from its latch against the direction of edges in CFG. The loops are organized in a containment hierarchy (tree) such that all the loops immediately contained inside loop L are the children of L in the tree. This tree is represented by the struct loops structure. The root of this tree is a fake loop that contains all blocks in the function. Each of the loops is represented in a struct loop structure. Each loop is assigned an index (num field of the struct loop structure), and the pointer to the loop is stored in the corresponding field of the larray vector in the loops structure. The indices do not have to be continuous, there may be empty (NULL) entries in the larray created by deleting loops. Also, there is no guarantee on the relative order of a loop and its subloops in the numbering. The index of a loop never changes.

The entries of the larray field should not be accessed directly. The function get_loop returns the loop description for a loop with the given index. number_of_loops function returns number of loops in the function. To traverse all loops, use FOR_EACH_LOOP macro. The flags argument of the macro is used to determine the direction of traversal and the set of loops visited. Each loop is guaranteed to be visited exactly once, regardless of the changes to the loop tree, and the loops may be removed during the traversal. The newly created loops are never traversed, if they need to be visited, this must be done separately after their creation. The FOR_EACH_LOOP macro allocates temporary variables. If the FOR_EACH_LOOP
loop were ended using break or goto, they would not be released; `FOR_EACH_LOOP_BREAK` macro must be used instead.

Each basic block contains the reference to the innermost loop it belongs to (`loop_father`). For this reason, it is only possible to have one `struct loops` structure initialized at the same time for each CFG. The global variable `current_loops` contains the `struct loops` structure. Many of the loop manipulation functions assume that dominance information is up-to-date.

The loops are analyzed through `loop_optimizer_init` function. The argument of this function is a set of flags represented in an integer bitmask. These flags specify what other properties of the loop structures should be calculated/enforced and preserved later:

- **LOOPS_MAY_HAVE_MULTIPLE_LATCHES**: If this flag is set, no changes to CFG will be performed in the loop analysis, in particular, loops with multiple latch edges will not be disambiguated. If a loop has multiple latches, its latch block is set to NULL. Most of the loop manipulation functions will not work for loops in this shape. No other flags that require CFG changes can be passed to `loop_optimizer_init`.

- **LOOPS_HAVE_PREHEADERS**: Forwarder blocks are created in such a way that each loop has only one entry edge, and additionally, the source block of this entry edge has only one successor. This creates a natural place where the code can be moved out of the loop, and ensures that the entry edge of the loop leads from its immediate super-loop.

- **LOOPS_HAVE_SIMPLE_LATCHES**: Forwarder blocks are created to force the latch block of each loop to have only one successor. This ensures that the latch of the loop does not belong to any of its sub-loops, and makes manipulation with the loops significantly easier. Most of the loop manipulation functions assume that the loops are in this shape. Note that with this flag, the “normal” loop without any control flow inside and with one exit consists of two basic blocks.

- **LOOPS_HAVE_MARKED_IRREDUCIBLE_REGIONS**: Basic blocks and edges in the strongly connected components that are not natural loops (have more than one entry block) are marked with `BB_IRREDUCIBLE_LOOP` and `EDGE_IRREDUCIBLE_LOOP` flags. The flag is not set for blocks and edges that belong to natural loops that are in such an irreducible region (but it is set for the entry and exit edges of such a loop, if they lead to/from this region).

- **LOOPS_HAVE_RECORDED_EXITS**: The lists of exits are recorded and updated for each loop. This makes some functions (e.g., `get_loop_exit_edges`) more efficient. Some functions (e.g., `single_exit`) can be used only if the lists of exits are recorded.

These properties may also be computed/enforced later, using functions `create_preheaders`, `force_single_succ_latches`, `mark_irreducible_loops` and `record_loop_exits`.

The memory occupied by the loops structures should be freed with `loop_optimizer_finalize` function.

The CFG manipulation functions in general do not update loop structures. Specialized versions that additionally do so are provided for the most common tasks. On GIMPLE, `cleanup_tree_cfg_loop` function can be used to cleanup CFG while updating the loops structures if `current_loops` is set.
14.2 Loop querying

The functions to query the information about loops are declared in ‘cfgloop.h’. Some of the information can be taken directly from the structures. `loop_father` field of each basic block contains the innermost loop to that the block belongs. The most useful fields of loop structure (that are kept up-to-date at all times) are:

- **header, latch**: Header and latch basic blocks of the loop.
- **num_nodes**: Number of basic blocks in the loop (including the basic blocks of the sub-loops).
- **depth**: The depth of the loop in the loops tree, i.e., the number of super-loops of the loop.
- **outer, inner, next**: The super-loop, the first sub-loop, and the sibling of the loop in the loops tree.

There are other fields in the loop structures, many of them used only by some of the passes, or not updated during CFG changes; in general, they should not be accessed directly.

The most important functions to query loop structures are:

- **flow_loops_dump**: Dumps the information about loops to a file.
- **verify_loop_structure**: Checks consistency of the loop structures.
- **loop_latch_edge**: Returns the latch edge of a loop.
- **loop_preheader_edge**: Returns the preheader edge of a loop.
- **flow_loop_nested_p**: Tests whether loop is a sub-loop of another loop.
- **flow_bb_inside_loop_p**: Tests whether a basic block belongs to a loop (including its sub-loops).
- **find_common_loop**: Finds the common super-loop of two loops.
- **superloop_at_depth**: Returns the super-loop of a loop with the given depth.
- **tree_num_loop_insns, num_loop_insns**: Estimates the number of insns in the loop, on GIMPLE and on RTL.
- **loop_exit_edge_p**: Tests whether edge is an exit from a loop.
- **mark_loop_exit_edges**: Marks all exit edges of all loops with EDGE_LOOP_EXIT flag.
- **get_loop_body, get_loop_body_in_dom_order, get_loop_body_in_bfs_order**: Enumerates the basic blocks in the loop in depth-first search order in reversed CFG, ordered by dominance relation, and breadth-first search order, respectively.
- **single_exit**: Returns the single exit edge of the loop, or NULL if the loop has more than one exit. You can only use this function if LOOPS_HAVE_MARKED_SINGLE_EXITS property is used.
- **get_loop_exit_edges**: Enumerates the exit edges of a loop.
- **just_once_each_iteration_p**: Returns true if the basic block is executed exactly once during each iteration of a loop (that is, it does not belong to a sub-loop, and it dominates the latch of the loop).
14.3 Loop manipulation

The loops tree can be manipulated using the following functions:

- **flow_loop_tree_node_add**: Adds a node to the tree.
- **flow_loop_tree_node_remove**: Removes a node from the tree.
- **add_bb_to_loop**: Adds a basic block to a loop.
- **remove_bb_from_loops**: Removes a basic block from loops.

Most low-level CFG functions update loops automatically. The following functions handle some more complicated cases of CFG manipulations:

- **remove_path**: Removes an edge and all blocks it dominates.
- **split_loop_exit_edge**: Splits exit edge of the loop, ensuring that PHI node arguments remain in the loop (this ensures that loop-closed SSA form is preserved). Only useful on GIMPLE.

Finally, there are some higher-level loop transformations implemented. While some of them are written so that they should work on non-innermost loops, they are mostly untested in that case, and at the moment, they are only reliable for the innermost loops:

- **create_iv**: Creates a new induction variable. Only works on GIMPLE. **standard_iv_increment_position** can be used to find a suitable place for the iv increment.
- **duplicate_loop_to_header_edge, tree_duplicate_loop_to_header_edge**: These functions (on RTL and on GIMPLE) duplicate the body of the loop prescribed number of times on one of the edges entering loop header, thus performing either loop unrolling or loop peeling. **can_duplicate_loop_p** (**can_unroll_loop_p** on GIMPLE) must be true for the duplicated loop.
- **loop_version, tree_ssa_loop_version**: These function create a copy of a loop, and a branch before them that selects one of them depending on the prescribed condition. This is useful for optimizations that need to verify some assumptions in runtime (one of the copies of the loop is usually left unchanged, while the other one is transformed in some way).
- **tree_unroll_loop**: Unrolls the loop, including peeling the extra iterations to make the number of iterations divisible by unroll factor, updating the exit condition, and removing the exits that now cannot be taken. Works only on GIMPLE.

14.4 Loop-closed SSA form

Throughout the loop optimizations on tree level, one extra condition is enforced on the SSA form: No SSA name is used outside of the loop in that it is defined. The SSA form satisfying this condition is called “loop-closed SSA form” – LCSSA. To enforce LCSSA, PHI nodes must be created at the exits of the loops for the SSA names that are used outside of them. Only the real operands (not virtual SSA names) are held in LCSSA, in order to save memory.

There are various benefits of LCSSA:

- Many optimizations (value range analysis, final value replacement) are interested in the values that are defined in the loop and used outside of it, i.e., exactly those for that we create new PHI nodes.
In induction variable analysis, it is not necessary to specify the loop in that the analysis should be performed – the scalar evolution analysis always returns the results with respect to the loop in that the SSA name is defined.

- It makes updating of SSA form during loop transformations simpler. Without LCSSA, operations like loop unrolling may force creation of PHI nodes arbitrarily far from the loop, while in LCSSA, the SSA form can be updated locally. However, since we only keep real operands in LCSSA, we cannot use this advantage (we could have local updating of real operands, but it is not much more efficient than to use generic SSA form updating for it as well; the amount of changes to SSA is the same).

    However, it also means LCSSA must be updated. This is usually straightforward, unless you create a new value in loop and use it outside, or unless you manipulate loop exit edges (functions are provided to make these manipulations simple). rewrite_into_loop_closed_ssa is used to rewrite SSA form to LCSSA, and verify_loop_closed_ssa to check that the invariant of LCSSA is preserved.

14.5 Scalar evolutions

Scalar evolutions (SCEV) are used to represent results of induction variable analysis on GIMPLE. They enable us to represent variables with complicated behavior in a simple and consistent way (we only use it to express values of polynomial induction variables, but it is possible to extend it). The interfaces to SCEV analysis are declared in 'tree-scalar-evolution.h'. To use scalar evolutions analysis, scev_initialize must be used. To stop using SCEV, scev_finalize should be used. SCEV analysis caches results in order to save time and memory. This cache however is made invalid by most of the loop transformations, including removal of code. If such a transformation is performed, scev_reset must be called to clean the caches.

Given an SSA name, its behavior in loops can be analyzed using the analyze_scalar_evolution function. The returned SCEV however does not have to be fully analyzed and it may contain references to other SSA names defined in the loop. To resolve these (potentially recursive) references, instantiate_parameters or resolve_mixers functions must be used. instantiate_parameters is useful when you use the results of SCEV only for some analysis, and when you work with whole nest of loops at once. It will try replacing all SSA names by their SCEV in all loops, including the super-loops of the current loop, thus providing a complete information about the behavior of the variable in the loop nest. resolve_mixers is useful if you work with only one loop at a time, and if you possibly need to create code based on the value of the induction variable. It will only resolve the SSA names defined in the current loop, leaving the SSA names defined outside unchanged, even if their evolution in the outer loops is known.

The SCEV is a normal tree expression, except for the fact that it may contain several special tree nodes. One of them is SCEV_NOT_KNOWN, used for SSA names whose value cannot be expressed. The other one is POLYNOMIAL_CHREC. Polynomial chrec has three arguments – base, step and loop (both base and step may contain further polynomial chrecs). Type of the expression and of base and step must be the same. A variable has evolution POLYNOMIAL_CHREC(base, step, loop) if it is (in the specified loop) equivalent to x_1 in the following example

    while (...)
\[
\{ 
    x_1 = \text{phi} (\text{base}, x_2);
    x_2 = x_1 + \text{step};
\}
\]

Note that this includes the language restrictions on the operations. For example, if we compile C code and \( x \) has signed type, then the overflow in addition would cause undefined behavior, and we may assume that this does not happen. Hence, the value with this SCEV cannot overflow (which restricts the number of iterations of such a loop).

In many cases, one wants to restrict the attention just to affine induction variables. In this case, the extra expressive power of SCEV is not useful, and may complicate the optimizations. In this case, \texttt{simple_iv} function may be used to analyze a value – the result is a loop-invariant base and step.

### 14.6 IV analysis on RTL

The induction variable on RTL is simple and only allows analysis of affine induction variables, and only in one loop at once. The interface is declared in \texttt{cfgloop.h}. Before analyzing induction variables in a loop \( L \), \texttt{iv_analysis_loop_init} function must be called on \( L \). After the analysis (possibly calling \texttt{iv_analysis_loop_init} for several loops) is finished, \texttt{iv_analysis_done} should be called. The following functions can be used to access the results of the analysis:

- \texttt{iv_analyze}: Analyzes a single register used in the given insn. If no use of the register in this insn is found, the following insns are scanned, so that this function can be called on the insn returned by get_condition.
- \texttt{iv_analyze_result}: Analyzes result of the assignment in the given insn.
- \texttt{iv_analyze_expr}: Analyzes a more complicated expression. All its operands are analyzed by \texttt{iv_analyze}, and hence they must be used in the specified insn or one of the following insns.

The description of the induction variable is provided in \texttt{struct rtx_iv}. In order to handle subregs, the representation is a bit complicated; if the value of the \texttt{extend} field is not \texttt{UNKNOWN}, the value of the induction variable in the \( i \)-th iteration is

\[
\delta + \text{mult} \times \text{extend}_{\text{extend_mode}}(\text{subreg}_{\text{mode}}(\text{base} + i \times \text{step})),
\]

with the following exception: if \texttt{first_special} is true, then the value in the first iteration (when \( i \) is zero) is \texttt{delta + mult * base}. However, if \texttt{extend} is equal to \texttt{UNKNOWN}, then \texttt{first_special} must be false, \texttt{delta 0, mult 1} and the value in the \( i \)-th iteration is

\[
\text{subreg}_{\text{mode}}(\text{base} + i \times \text{step})
\]

The function \texttt{get_iv_value} can be used to perform these calculations.

### 14.7 Number of iterations analysis

Both on GIMPLE and on RTL, there are functions available to determine the number of iterations of a loop, with a similar interface. The number of iterations of a loop in GCC is defined as the number of executions of the loop latch. In many cases, it is not possible to determine the number of iterations unconditionally – the determined number is correct only if some assumptions are satisfied. The analysis tries to verify these conditions using the information contained in the program; if it fails, the conditions are returned together with the result. The following information and conditions are provided by the analysis:
• **assumptions**: If this condition is false, the rest of the information is invalid.

• **noloop_assumptions on RTL, may_be_zero on GIMPLE**: If this condition is true, the loop exits in the first iteration.

• **infinite**: If this condition is true, the loop is infinite. This condition is only available on RTL. On GIMPLE, conditions for finiteness of the loop are included in assumptions.

• **niter_expr on RTL, niter on GIMPLE**: The expression that gives number of iterations. The number of iterations is defined as the number of executions of the loop latch.

Both on GIMPLE and on RTL, it necessary for the induction variable analysis framework to be initialized (SCEV on GIMPLE, loop-iv on RTL). On GIMPLE, the results are stored to `struct tree_niter_desc` structure. Number of iterations before the loop is exited through a given exit can be determined using `number_of_iterations_exit` function. On RTL, the results are returned in `struct niter_desc` structure. The corresponding function is named `check_simple_exit`. There are also functions that pass through all the exits of a loop and try to find one with easy to determine number of iterations – `find_loop_niter` on GIMPLE and `find_simple_exit` on RTL. Finally, there are functions that provide the same information, but additionally cache it, so that repeated calls to number of iterations are not so costly – `number_of_latch_executions` on GIMPLE and `get_simple_loop_desc` on RTL.

Note that some of these functions may behave slightly differently than others – some of them return only the expression for the number of iterations, and fail if there are some assumptions. The function `number_of_latch_executions` works only for single-exit loops. The function `number_of_cond_exit_executions` can be used to determine number of executions of the exit condition of a single-exit loop (i.e., the `number_of_latch_executions` increased by one).

### 14.8 Data Dependency Analysis

The code for the data dependence analysis can be found in `tree-data-ref.c` and its interface and data structures are described in `tree-data-ref.h`. The function that computes the data dependences for all the array and pointer references for a given loop is `compute_data_dependences_for_loop`. This function is currently used by the linear loop transform and the vectorization passes. Before calling this function, one has to allocate two vectors: a first vector will contain the set of data references that are contained in the analyzed loop body, and the second vector will contain the dependence relations between the data references. Thus if the vector of data references is of size `n`, the vector containing the dependence relations will contain `n*n` elements. However if the analyzed loop contains side effects, such as calls that potentially can interfere with the data references in the current analyzed loop, the analysis stops while scanning the loop body for data references, and inserts a single `chrec_dont_know` in the dependence relation array.

The data references are discovered in a particular order during the scanning of the loop body: the loop body is analyzed in execution order, and the data references of each statement are pushed at the end of the data reference array. Two data references syntactically occur in the program in the same order as in the array of data references. This syntactic
order is important in some classical data dependence tests, and mapping this order to the elements of this array avoids costly queries to the loop body representation.

Three types of data references are currently handled: ARRAY_REF, INDIRECT_REF and COMPONENT_REF. The data structure for the data reference is data_reference, where data_reference_p is a name of a pointer to the data reference structure. The structure contains the following elements:

- **base_object_info**: Provides information about the base object of the data reference and its access functions. These access functions represent the evolution of the data reference in the loop relative to its base, in keeping with the classical meaning of the data reference access function for the support of arrays. For example, for a reference a.b[i][j], the base object is a.b and the access functions, one for each array subscript, are: {i_init, + i_step}_1, {j_init, +, j_step}_2.

- **first_location_in_loop**: Provides information about the first location accessed by the data reference in the loop and about the access function used to represent evolution relative to this location. This data is used to support pointers, and is not used for arrays (for which we have base objects). Pointer accesses are represented as a one-dimensional access that starts from the first location accessed in the loop. For example:

```c
for1 i
  for2 j
    *((int *)p + i + j) = a[i][j];
```

The access function of the pointer access is {0, + 4B}_for2 relative to p + i. The access functions of the array are {i_init, + i_step}_for1 and {j_init, +, j_step}_for2 relative to a.

Usually, the object the pointer refers to is either unknown, or we can't prove that the access is confined to the boundaries of a certain object.

Two data references can be compared only if at least one of these two representations has all its fields filled for both data references.

The current strategy for data dependence tests is as follows: If both a and b are represented as arrays, compare a.base_object and b.base_object; if they are equal, apply dependence tests (use access functions based on base_objects). Else if both a and b are represented as pointers, compare a.first_location and b.first_location; if they are equal, apply dependence tests (use access functions based on first location). However, if a and b are represented differently, only try to prove that the bases are definitely different.

- Aliasing information.
- Alignment information.

The structure describing the relation between two data references is data_dependence_relation and the shorter name for a pointer to such a structure is ddr_p. This structure contains:

- a pointer to each data reference,
- a tree node are_dependent that is set to chrec_known if the analysis has proved that there is no dependence between these two data references, chrec_dont_know if the analysis was not able to determine any useful result and potentially there could exist a dependence between these data references, and are_dependent is set to NULL_TREE
if there exist a dependence relation between the data references, and the description of this dependence relation is given in the subscripts, dir_vecs, and dist_vecs arrays,

- a boolean that determines whether the dependence relation can be represented by a classical distance vector,
- an array subscripts that contains a description of each subscript of the data references. Given two array accesses a subscript is the tuple composed of the access functions for a given dimension. For example, given \( A[f_1][f_2][f_3] \) and \( B[g_1][g_2][g_3] \), there are three subscripts: \((f_1, g_1), (f_2, g_2), (f_3, g_3)\).
- two arrays dir_vecs and dist_vecs that contain classical representations of the data dependences under the form of direction and distance dependence vectors,
- an array of loops loop_nest that contains the loops to which the distance and direction vectors refer to.

Several functions for pretty printing the information extracted by the data dependence analysis are available: dump_ddrs prints with a maximum verbosity the details of a data dependence relations array, dump_dist_dir_vectors prints only the classical distance and direction vectors for a data dependence relations array, and dump_data_references prints the details of the data references contained in a data reference array.

### 14.9 Linear loop transformations framework

Lambda is a framework that allows transformations of loops using non-singular matrix based transformations of the iteration space and loop bounds. This allows compositions of skewing, scaling, interchange, and reversal transformations. These transformations are often used to improve cache behavior or remove inner loop dependencies to allow parallelization and vectorization to take place.

To perform these transformations, Lambda requires that the loopnest be converted into an internal form that can be matrix transformed easily. To do this conversion, the function gcc_loopnest_to_lambda_loopnest is provided. If the loop cannot be transformed using lambda, this function will return NULL.

Once a lambda_loopnest is obtained from the conversion function, it can be transformed by using lambda_loopnest_transform, which takes a transformation matrix to apply. Note that it is up to the caller to verify that the transformation matrix is legal to apply to the loop (dependence respecting, etc). Lambda simply applies whatever matrix it is told to provide. It can be extended to make legal matrices out of any non-singular matrix, but this is not currently implemented. Legality of a matrix for a given loopnest can be verified using lambda_transform_legal_p.

Given a transformed loopnest, conversion back into gcc IR is done by lambda_loopnest_to_gcc_loopnest. This function will modify the loops so that they match the transformed loopnest.

### 14.10 Omega a solver for linear programming problems

The data dependence analysis contains several solvers triggered sequentially from the less complex ones to the more sophisticated. For ensuring the consistency of the results of these solvers, a data dependence check pass has been implemented based on two different solvers.
The second method that has been integrated to GCC is based on the Omega dependence solver, written in the 1990’s by William Pugh and David Wonnacott. Data dependence tests can be formulated using a subset of the Presburger arithmetics that can be translated to linear constraint systems. These linear constraint systems can then be solved using the Omega solver.

The Omega solver is using Fourier-Motzkin’s algorithm for variable elimination: a linear constraint system containing \( n \) variables is reduced to a linear constraint system with \( n-1 \) variables. The Omega solver can also be used for solving other problems that can be expressed under the form of a system of linear equalities and inequalities. The Omega solver is known to have an exponential worst case, also known under the name of “omega nightmare” in the literature, but in practice, the omega test is known to be efficient for the common data dependence tests.

The interface used by the Omega solver for describing the linear programming problems is described in `omega.h`, and the solver is `omega_solve_problem`. 
15 Control Flow Graph

A control flow graph (CFG) is a data structure built on top of the intermediate code representation (the RTL or tree instruction stream) abstracting the control flow behavior of a function that is being compiled. The CFG is a directed graph where the vertices represent basic blocks and edges represent possible transfer of control flow from one basic block to another. The data structures used to represent the control flow graph are defined in ‘basic-block.h’.

15.1 Basic Blocks

A basic block is a straight-line sequence of code with only one entry point and only one exit. In GCC, basic blocks are represented using the basic_block data type.

Two pointer members of the basic_block structure are the pointers next_bb and prev_bb. These are used to keep doubly linked chain of basic blocks in the same order as the underlying instruction stream. The chain of basic blocks is updated transparently by the provided API for manipulating the CFG. The macro FOR_EACH_BB can be used to visit all the basic blocks in lexicographical order. Dominator traversals are also possible using walk_dominator_tree. Given two basic blocks A and B, block A dominates block B if A is always executed before B.

The BASIC_BLOCK array contains all basic blocks in an unspecified order. Each basic_block structure has a field that holds a unique integer identifier index that is the index of the block in the BASIC_BLOCK array. The total number of basic blocks in the function is n_basic_blocks. Both the basic block indices and the total number of basic blocks may vary during the compilation process, as passes reorder, create, duplicate, and destroy basic blocks. The index for any block should never be greater than last_basic_block.

Special basic blocks represent possible entry and exit points of a function. These blocks are called ENTRY_BLOCK_PTR and EXIT_BLOCK_PTR. These blocks do not contain any code, and are not elements of the BASIC_BLOCK array. Therefore they have been assigned unique, negative index numbers.

Each basic_block also contains pointers to the first instruction (the head) and the last instruction (the tail) or end of the instruction stream contained in a basic block. In fact, since the basic_block data type is used to represent blocks in both major intermediate representations of GCC (tree and RTL), there are pointers to the head and end of a basic block for both representations.

For RTL, these pointers are rtx head, end. In the RTL function representation, the head pointer always points either to a NOTE_INSN_BASIC_BLOCK or to a CODE_LABEL, if present. In the RTL representation of a function, the instruction stream contains not only the “real” instructions, but also notes. Any function that moves or duplicates the basic blocks needs to take care of updating of these notes. Many of these notes expect that the instruction stream consists of linear regions, making such updates difficult. The NOTE_INSN_BASIC_BLOCK note is the only kind of note that may appear in the instruction stream contained in a basic block. The instruction stream of a basic block always follows a NOTE_INSN_BASIC_BLOCK, but zero or more CODE_LABEL nodes can precede the block note. A basic block ends by control flow instruction or last instruction before following CODE_LABEL or NOTE_INSN_BASIC_BLOCK. A CODE_LABEL cannot appear in the instruction stream of a basic block.
In addition to notes, the jump table vectors are also represented as “pseudo-instructions” inside the insn stream. These vectors never appear in the basic block and should always be placed just after the table jump instructions referencing them. After removing the table-jump it is often difficult to eliminate the code computing the address and referencing the vector, so cleaning up these vectors is postponed until after liveness analysis. Thus the jump table vectors may appear in the insn stream unreferenced and without any purpose. Before any edge is made fall-thru, the existence of such construct in the way needs to be checked by calling `can_fallthru` function.

For the `tree` representation, the head and end of the basic block are being pointed to by the `stmt_list` field, but this special `tree` should never be referenced directly. Instead, at the tree level abstract containers and iterators are used to access statements and expressions in basic blocks. These iterators are called `block statement iterators` (BSIs). Grep for `bsi` in the various ‘tree-*’ files. The following snippet will pretty-print all the statements of the program in the GIMPLE representation.

```c
FOR_EACH_BB (bb)
{
    block_stmt_iterator si;
    for (si = bsi_start (bb); !bsi_end_p (si); bsi_next (&si))
    {
        tree stmt = bsi_stmt (si);
        print_generic_stmt (stderr, stmt, 0);
    }
}
```

### 15.2 Edges

Edges represent possible control flow transfers from the end of some basic block A to the head of another basic block B. We say that A is a predecessor of B, and B is a successor of A. Edges are represented in GCC with the `edge` data type. Each `edge` acts as a link between two basic blocks: the `src` member of an edge points to the predecessor basic block of the `dest` basic block. The members `preds` and `succs` of the `basic_block` data type point to type-safe vectors of edges to the predecessors and successors of the block.

When walking the edges in an edge vector, `edge iterators` should be used. Edge iterators are constructed using the `edge_iterator` data structure and several methods are available to operate on them:

- **ei_start**: This function initializes an `edge_iterator` that points to the first edge in a vector of edges.
- **ei_last**: This function initializes an `edge_iterator` that points to the last edge in a vector of edges.
- **ei_end_p**: This predicate is `true` if an `edge_iterator` represents the last edge in an edge vector.
- **ei_one_before_end_p**: This predicate is `true` if an `edge_iterator` represents the second last edge in an edge vector.
- **ei_next**: This function takes a pointer to an `edge_iterator` and makes it point to the next edge in the sequence.
ei_prev   This function takes a pointer to an edge_iterator and makes it point to the previous edge in the sequence.

ei_edge   This function returns the edge currently pointed to by an edge_iterator.

ei_safe_safe   This function returns the edge currently pointed to by an edge_iterator, but returns NULL if the iterator is pointing at the end of the sequence. This function has been provided for existing code makes the assumption that a NULL edge indicates the end of the sequence.

The convenience macro FOR_EACH_EDGE can be used to visit all of the edges in a sequence of predecessor or successor edges. It must not be used when an element might be removed during the traversal, otherwise elements will be missed. Here is an example of how to use the macro:

```c
edge e;
edge_iterator ei;

FOR_EACH_EDGE (e, ei, bb->succs)
{
    if (e->flags & EDGE_FALLTHRU)
        break;
}
```

There are various reasons why control flow may transfer from one block to another. One possibility is that some instruction, for example a CODE_LABEL, in a linearized instruction stream just always starts a new basic block. In this case a fall-thru edge links the basic block to the first following basic block. But there are several other reasons why edges may be created. The flags field of the edge data type is used to store information about the type of edge we are dealing with. Each edge is of one of the following types:

**jump**   No type flags are set for edges corresponding to jump instructions. These edges are used for unconditional or conditional jumps and in RTL also for table jumps. They are the easiest to manipulate as they may be freely redirected when the flow graph is not in SSA form.

**fall-thru**   Fall-thru edges are present in case where the basic block may continue execution to the following one without branching. These edges have the EDGE_FALLTHRU flag set. Unlike other types of edges, these edges must come into the basic block immediately following in the instruction stream. The function force_nonfallthru is available to insert an unconditional jump in the case that redirection is needed. Note that this may require creation of a new basic block.

**exception handling**   Exception handling edges represent possible control transfers from a trapping instruction to an exception handler. The definition of “trapping” varies. In C++, only function calls can throw, but for Java, exceptions like division by zero or segmentation fault are defined and thus each instruction possibly throwing this kind of exception needs to be handled as control flow instruction. Exception edges have the EDGE_ABNORMAL and EDGE_EH flags set.
When updating the instruction stream it is easy to change possibly trapping instruction to non-trapping, by simply removing the exception edge. The opposite conversion is difficult, but should not happen anyway. The edges can be eliminated via `purge_dead_edges` call.

In the RTL representation, the destination of an exception edge is specified by `REG_EH_REGION` note attached to the insn. In case of a trapping call the `EDGE_ABNORMAL_CALL` flag is set too. In the `tree` representation, this extra flag is not set.

In the RTL representation, the predicate `may_trap_p` may be used to check whether instruction still may trap or not. For the tree representation, the `tree_could_trap_p` predicate is available, but this predicate only checks for possible memory traps, as in dereferencing an invalid pointer location.

**Sibling calls**

Sibling calls or tail calls terminate the function in a non-standard way and thus an edge to the exit must be present. `EDGE_SIBCALL` and `EDGE_ABNORMAL` are set in such case. These edges only exist in the RTL representation.

**Computed jumps**

Computed jumps contain edges to all labels in the function referenced from the code. All those edges have `EDGE_ABNORMAL` flag set. The edges used to represent computed jumps often cause compile time performance problems, since functions consisting of many taken labels and many computed jumps may have very dense flow graphs, so these edges need to be handled with special care. During the earlier stages of the compilation process, GCC tries to avoid such dense flow graphs by factoring computed jumps. For example, given the following series of jumps,

```c
goto *x;
[ ... ]
```

```c
goto *x;
[ ... ]
```

```c
goto *x;
[ ... ]
```

factoring the computed jumps results in the following code sequence which has a much simpler flow graph:

```c
goto y;
[ ... ]
```

```c
goto y;
[ ... ]
```

```c
goto y;
[ ... ]
```

```c
y:
goto *x;
```

However, the classic problem with this transformation is that it has a runtime cost in there resulting code: An extra jump. Therefore, the computed jumps are un-factored in the later passes of the compiler. Be aware of that when
you work on passes in that area. There have been numerous examples already where the compile time for code with unfactored computed jumps caused some serious headaches.

**nonlocal goto handlers**

GCC allows nested functions to return into caller using a `goto` to a label passed to as an argument to the callee. The labels passed to nested functions contain special code to cleanup after function call. Such sections of code are referred to as “nonlocal goto receivers”. If a function contains such nonlocal goto receivers, an edge from the call to the label is created with the `EDGE_ABNORMAL` and `EDGE_ABNORMAL_CALL` flags set.

**function entry points**

By definition, execution of function starts at basic block 0, so there is always an edge from the `ENTRY_BLOCK_PTR` to basic block 0. There is no `tree` representation for alternate entry points at this moment. In RTL, alternate entry points are specified by `CODE_LABEL` with `LABEL_ALTERNATE_NAME` defined. This feature is currently used for multiple entry point prologues and is limited to post-reload passes only. This can be used by back-ends to emit alternate prologues for functions called from different contexts. In future full support for multiple entry functions defined by Fortran 90 needs to be implemented.

**function exits**

In the pre-reload representation a function terminates after the last instruction in the insn chain and no explicit return instructions are used. This corresponds to the fall-thru edge into exit block. After reload, optimal RTL epilogues are used that use explicit (conditional) return instructions that are represented by edges with no flags set.

### 15.3 Profile information

In many cases a compiler must make a choice whether to trade speed in one part of code for speed in another, or to trade code size for code speed. In such cases it is useful to know information about how often some given block will be executed. That is the purpose for maintaining profile within the flow graph. GCC can handle profile information obtained through *profile feedback*, but it can also estimate branch probabilities based on statics and heuristics.

The feedback based profile is produced by compiling the program with instrumentation, executing it on a train run and reading the numbers of executions of basic blocks and edges back to the compiler while re-compiling the program to produce the final executable. This method provides very accurate information about where a program spends most of its time on the train run. Whether it matches the average run of course depends on the choice of train data set, but several studies have shown that the behavior of a program usually changes just marginally over different data sets.

When profile feedback is not available, the compiler may be asked to attempt to predict the behavior of each branch in the program using a set of heuristics (see `predict.def` for details) and compute estimated frequencies of each basic block by propagating the probabilities over the graph.
Each basic_block contains two integer fields to represent profile information: frequency and count. The frequency is an estimation how often is basic block executed within a function. It is represented as an integer scaled in the range from 0 to BB_FREQ_BASE. The most frequently executed basic block in function is initially set to BB_FREQ_BASE and the rest of frequencies are scaled accordingly. During optimization, the frequency of the most frequent basic block can both decrease (for instance by loop unrolling) or grow (for instance by cross-jumping optimization), so scaling sometimes has to be performed multiple times.

The count contains hard-counted numbers of execution measured during training runs and is nonzero only when profile feedback is available. This value is represented as the host’s widest integer (typically a 64 bit integer) of the special type gcov_type.

Most optimization passes can use only the frequency information of a basic block, but a few passes may want to know hard execution counts. The frequencies should always match the counts after scaling, however during updating of the profile information numerical error may accumulate into quite large errors.

Each edge also contains a branch probability field: an integer in the range from 0 to REG_BR_PROB_BASE. It represents probability of passing control from the end of the src basic block to the dest basic block, i.e. the probability that control will flow along this edge. The EDGE_FREQUENCY macro is available to compute how frequently a given edge is taken. There is a count field for each edge as well, representing same information as for a basic block.

The basic block frequencies are not represented in the instruction stream, but in the RTL representation the edge frequencies are represented for conditional jumps (via the REG_BR_PROB macro) since they are used when instructions are output to the assembly file and the flow graph is no longer maintained.

The probability that control flow arrives via a given edge to its destination basic block is called reverse probability and is not directly represented, but it may be easily computed from frequencies of basic blocks.

Updating profile information is a delicate task that can unfortunately not be easily integrated with the CFG manipulation API. Many of the functions and hooks to modify the CFG, such as redirect_edge_and_branch, do not have enough information to easily update the profile, so updating it is in the majority of cases left up to the caller. It is difficult to uncover bugs in the profile updating code, because they manifest themselves only by producing worse code, and checking profile consistency is not possible because of numeric error accumulation. Hence special attention needs to be given to this issue in each pass that modifies the CFG.

It is important to point out that REG_BR_PROB_BASE and BB_FREQ_BASE are both set low enough to be possible to compute second power of any frequency or probability in the flow graph, it is not possible to even square the count field, as modern CPUs are fast enough to execute $2^{32}$ operations quickly.

15.4 Maintaining the CFG

An important task of each compiler pass is to keep both the control flow graph and all profile information up-to-date. Reconstruction of the control flow graph after each pass is not an option, since it may be very expensive and lost profile information cannot be reconstructed at all.
GCC has two major intermediate representations, and both use the `basic_block` and `edge` data types to represent control flow. Both representations share as much of the CFG maintenance code as possible. For each representation, a set of hooks is defined so that each representation can provide its own implementation of CFG manipulation routines when necessary. These hooks are defined in ‘`cfghooks.h`’. There are hooks for almost all common CFG manipulations, including block splitting and merging, edge redirection and creating and deleting basic blocks. These hooks should provide everything you need to maintain and manipulate the CFG in both the RTL and tree representation.

At the moment, the basic block boundaries are maintained transparently when modifying instructions, so there rarely is a need to move them manually (such as in case someone wants to output instruction outside basic block explicitly). Often the CFG may be better viewed as integral part of instruction chain, than structure built on the top of it. However, in principle the control flow graph for the tree representation is not an integral part of the representation, in that a function tree may be expanded without first building a flow graph for the tree representation at all. This happens when compiling without any tree optimization enabled. When the tree optimizations are enabled and the instruction stream is rewritten in SSA form, the CFG is very tightly coupled with the instruction stream. In particular, statement insertion and removal has to be done with care. In fact, the whole tree representation can not be easily used or maintained without proper maintenance of the CFG simultaneously.

In the RTL representation, each instruction has a `BLOCK_FOR_INSN` value that represents pointer to the basic block that contains the instruction. In the tree representation, the function `bb_for_stmt` returns a pointer to the basic block containing the queried statement.

When changes need to be applied to a function in its tree representation, block statement iterators should be used. These iterators provide an integrated abstraction of the flow graph and the instruction stream. Block statement iterators are constructed using the `block_stmt_iterator` data structure and several modifier are available, including the following:

- **bsi_start**
  This function initializes a `block_stmt_iterator` that points to the first non-empty statement in a basic block.

- **bsi_last**
  This function initializes a `block_stmt_iterator` that points to the last statement in a basic block.

- **bsi_end_p**
  This predicate is true if a `block_stmt_iterator` represents the end of a basic block.

- **bsi_next**
  This function takes a `block_stmt_iterator` and makes it point to its successor.

- **bsi_prev**
  This function takes a `block_stmt_iterator` and makes it point to its predecessor.

- **bsi_insert_after**
  This function inserts a statement after the `block_stmt_iterator` passed in. The final parameter determines whether the statement iterator is updated to point to the newly inserted statement, or left pointing to the original statement.
bsi_insert_before
This function inserts a statement before the block_stmt_iterator passed in. The final parameter determines whether the statement iterator is updated to point to the newly inserted statement, or left pointing to the original statement.

bsi_remove
This function removes the block_stmt_iterator passed in and rechains the remaining statements in a basic block, if any.

In the RTL representation, the macros BB_HEAD and BB_END may be used to get the head and end rtx of a basic block. No abstract iterators are defined for traversing the insn chain, but you can just use NEXT_INSN and PREV_INSN instead. See Section 10.19 [Insns], page 166.

Usually a code manipulating pass simplifies the instruction stream and the flow of control, possibly eliminating some edges. This may for example happen when a conditional jump is replaced with an unconditional jump, but also when simplifying possibly trapping instruction to non-trapping while compiling Java. Updating of edges is not transparent and each optimization pass is required to do so manually. However only few cases occur in practice. The pass may call purge_dead_edges on a given basic block to remove superfluous edges, if any.

Another common scenario is redirection of branch instructions, but this is best modeled as redirection of edges in the control flow graph and thus use of redirect_edge_and_branch is preferred over more low level functions, such as redirect_jump that operate on RTL chain only. The CFG hooks defined in ‘cfghooks.h’ should provide the complete API required for manipulating and maintaining the CFG.

It is also possible that a pass has to insert control flow instruction into the middle of a basic block, thus creating an entry point in the middle of the basic block, which is impossible by definition: The block must be split to make sure it only has one entry point, i.e. the head of the basic block. The CFG hook split_block may be used when an instruction in the middle of a basic block has to become the target of a jump or branch instruction.

For a global optimizer, a common operation is to split edges in the flow graph and insert instructions on them. In the RTL representation, this can be easily done using the insert_insn_on_edge function that emits an instruction “on the edge”, caching it for a later commit_edge_insertions call that will take care of moving the inserted instructions off the edge into the instruction stream contained in a basic block. This includes the creation of new basic blocks where needed. In the tree representation, the equivalent functions are bsi_insert_on_edge which inserts a block statement iterator on an edge, and bsi_commit_edge_inserts which flushes the instruction to actual instruction stream.

While debugging the optimization pass, a verify_flow_info function may be useful to find bugs in the control flow graph updating code.

Note that at present, the representation of control flow in the tree representation is discarded before expanding to RTL. Long term the CFG should be maintained and “expanded” to the RTL representation along with the function tree itself.

15.5 Liveness information
Liveness information is useful to determine whether some register is “live” at given point of program, i.e. that it contains a value that may be used at a later point in the program.
This information is used, for instance, during register allocation, as the pseudo registers only need to be assigned to a unique hard register or to a stack slot if they are live. The hard registers and stack slots may be freely reused for other values when a register is dead.

Liveness information is available in the back end starting with pass_df_initialize and ending with pass_df_finish. Three flavors of live analysis are available: With LR, it is possible to determine at any point $P$ in the function if the register may be used on some path from $P$ to the end of the function. With UR, it is possible to determine if there is a path from the beginning of the function to $P$ that defines the variable. LIVE is the intersection of the LR and UR and a variable is live at $P$ if there is both an assignment that reaches it from the beginning of the function and a use that can be reached on some path from $P$ to the end of the function.

In general LIVE is the most useful of the three. The macros DF_[LR,UR,LIVE]_[IN,OUT] can be used to access this information. The macros take a basic block number and return a bitmap that is indexed by the register number. This information is only guaranteed to be up to date after calls are made to df_analyze. See the file df-core.c for details on using the dataflow.

The liveness information is stored partly in the RTL instruction stream and partly in the flow graph. Local information is stored in the instruction stream: Each instruction may contain REG_DEAD notes representing that the value of a given register is no longer needed, or REG_UNUSED notes representing that the value computed by the instruction is never used. The second is useful for instructions computing multiple values at once.
16 Machine Descriptions

A machine description has two parts: a file of instruction patterns (`.md` file) and a C header file of macro definitions.

The `.md` file for a target machine contains a pattern for each instruction that the target machine supports (or at least each instruction that is worth telling the compiler about). It may also contain comments. A semicolon causes the rest of the line to be a comment, unless the semicolon is inside a quoted string.

See the next chapter for information on the C header file.

16.1 Overview of How the Machine Description is Used

There are three main conversions that happen in the compiler:

1. The front end reads the source code and builds a parse tree.
2. The parse tree is used to generate an RTL insn list based on named instruction patterns.
3. The insn list is matched against the RTL templates to produce assembler code.

For the generate pass, only the names of the insns matter, from either a named `define_insn` or a `define_expand`. The compiler will choose the pattern with the right name and apply the operands according to the documentation later in this chapter, without regard for the RTL template or operand constraints. Note that the names the compiler looks for are hard-coded in the compiler—it will ignore unnamed patterns and patterns with names it doesn’t know about, but if you don’t provide a named pattern it needs, it will abort.

If a `define_insn` is used, the template given is inserted into the insn list. If a `define_expand` is used, one of three things happens, based on the condition logic. The condition logic may manually create new insns for the insn list, say via `emit_insn()`, and invoke `DONE`. For certain named patterns, it may invoke `FAIL` to tell the compiler to use an alternate way of performing that task. If it invokes neither `DONE` nor `FAIL`, the template given in the pattern is inserted, as if the `define_expand` were a `define_insn`.

Once the insn list is generated, various optimization passes convert, replace, and rearrange the insns in the insn list. This is where the `define_split` and `define_peephole` patterns get used, for example.

Finally, the insn list’s RTL is matched up with the RTL templates in the `define_insn` patterns, and those patterns are used to emit the final assembly code. For this purpose, each named `define_insn` acts like it’s unnamed, since the names are ignored.

16.2 Everything about Instruction Patterns

Each instruction pattern contains an incomplete RTL expression, with pieces to be filled in later, operand constraints that restrict how the pieces can be filled in, and an output pattern or C code to generate the assembler output, all wrapped up in a `define_insn` expression.

A `define_insn` is an RTL expression containing four or five operands:

1. An optional name. The presence of a name indicate that this instruction pattern can perform a certain standard job for the RTL-generation pass of the compiler. This pass knows certain names and will use the instruction patterns with those names, if the names are defined in the machine description.
The absence of a name is indicated by writing an empty string where the name should go. Nameless instruction patterns are never used for generating RTL code, but they may permit several simpler insns to be combined later on.

Names that are not thus known and used in RTL-generation have no effect; they are equivalent to no name at all.

For the purpose of debugging the compiler, you may also specify a name beginning with the '*' character. Such a name is used only for identifying the instruction in RTL dumps; it is entirely equivalent to having a nameless pattern for all other purposes.

2. The RTL template (see Section 16.4 [RTL Template], page 291) is a vector of incomplete RTL expressions which show what the instruction should look like. It is incomplete because it may contain match_operand, match_operator, and match_dup expressions that stand for operands of the instruction.

If the vector has only one element, that element is the template for the instruction pattern. If the vector has multiple elements, then the instruction pattern is a parallel expression containing the elements described.

3. A condition. This is a string which contains a C expression that is the final test to decide whether an insn body matches this pattern.

For a named pattern, the condition (if present) may not depend on the data in the insn being matched, but only the target-machine-type flags. The compiler needs to test these conditions during initialization in order to learn exactly which named instructions are available in a particular run.

For nameless patterns, the condition is applied only when matching an individual insn, and only after the insn has matched the pattern’s recognition template. The insn’s operands may be found in the vector operands. For an insn where the condition has once matched, it can’t be used to control register allocation, for example by excluding certain hard registers or hard register combinations.

4. The output template: a string that says how to output matching insns as assembler code. '%%' in this string specifies where to substitute the value of an operand. See Section 16.5 [Output Template], page 295.

When simple substitution isn’t general enough, you can specify a piece of C code to compute the output. See Section 16.6 [Output Statement], page 296.

5. Optionally, a vector containing the values of attributes for insns matching this pattern. See Section 16.19 [Insn Attributes], page 371.

### 16.3 Example of define_insn

Here is an actual example of an instruction pattern, for the 68000/68020.

```c
(define_insn "tstsi"
   [(set (cc0)
       (match_operand:SI 0 "general_operand" "rm"))]
   ""
   "*
{
   if (TARGET_68020 || !ADDRESS_REG_P (operands[0]))
     return \"tstl %0\";
   return \"cmpl #0,%0\";
}"
```
This can also be written using braced strings:

```
(define_insn "tstsi"
  [(set (cc0)
       (match_operand:SI 0 "general_operand" "rm"))
   ""
   {
   if (TARGET_68020 || ! ADDRESS_REG_P (operands[0]))
     return "tstl %0";
   return "cmpl #0,%0";
   })
```

This is an instruction that sets the condition codes based on the value of a general operand. It has no condition, so any insn whose RTL description has the form shown may be handled according to this pattern. The name ‘tstsi’ means “test a Simode value” and tells the RTL generation pass that, when it is necessary to test such a value, an insn to do so can be constructed using this pattern.

The output control string is a piece of C code which chooses which output template to return based on the kind of operand and the specific type of CPU for which code is being generated.

‘"rm"’ is an operand constraint. Its meaning is explained below.

### 16.4 RTL Template

The RTL template is used to define which insns match the particular pattern and how to find their operands. For named patterns, the RTL template also says how to construct an insn from specified operands.

Construction involves substituting specified operands into a copy of the template. Matching involves determining the values that serve as the operands in the insn being matched. Both of these activities are controlled by special expression types that direct matching and substitution of the operands.

```
(match_operand: m n predicate constraint)
```

This expression is a placeholder for operand number n of the insn. When constructing an insn, operand number n will be substituted at this point. When matching an insn, whatever appears at this position in the insn will be taken as operand number n; but it must satisfy predicate or this instruction pattern will not match at all.

Operand numbers must be chosen consecutively counting from zero in each instruction pattern. There may be only one match_operand expression in the pattern for each operand number. Usually operands are numbered in the order of appearance in match_operand expressions. In the case of a define_expand, any operand numbers used only in match_dup expressions have higher values than all other operand numbers.

predicate is a string that is the name of a function that accepts two arguments, an expression and a machine mode. See Section 16.7 [Predicates], page 297. During matching, the function will be called with the putative operand as the expression and m as the mode argument (if m is not specified, VOIDmode will be used, which normally causes predicate to accept any mode). If it returns zero, this instruction pattern fails to match. predicate may be an empty string; then
it means no test is to be done on the operand, so anything which occurs in this
position is valid.
Most of the time, predicate will reject modes other than \texttt{m}—but not always.
For example, the predicate \texttt{address\_operand} uses \texttt{m} as the mode of memory
ref that the address should be valid for. Many predicates accept \texttt{const\_int}
nodes even though their mode is \texttt{VOIDmode}.
\texttt{constraint} controls reloading and the choice of the best register class to use for
a value, as explained later (see Section 16.8 [Constraints], page 302). If the
constraint would be an empty string, it can be omitted.
People are often unclear on the difference between the constraint and the predic-
te. The predicate helps decide whether a given insn matches the pattern. The
constraint plays no role in this decision; instead, it controls various decisions in
the case of an insn which does match.

\texttt{(match\_scratch:mm n constraint)}
This expression is also a placeholder for operand number \texttt{n} and indicates that
operand must be a \texttt{scratch} or \texttt{reg} expression.
When matching patterns, this is equivalent to
\texttt{(match\_operand:mm n "scratch\_operand" pred)}
but, when generating RTL, it produces a \texttt{(scratch:mm)} expression.
If the last few expressions in a \texttt{parallel} are \texttt{clobber} expressions whose
operands are either a hard register or \texttt{match\_scratch}, the combiner can add
or delete them when necessary. See Section 10.15 [Side Effects], page 159.

\texttt{(match\_dup n)}
This expression is also a placeholder for operand number \texttt{n}. It is used when the
operand needs to appear more than once in the insn.
In construction, \texttt{match\_dup} acts just like \texttt{match\_operand}: the operand is sub-
stituted into the insn being constructed. But in matching, \texttt{match\_dup} behaves
differently. It assumes that operand number \texttt{n} has already been determined by
a \texttt{match\_operand} appearing earlier in the recognition template, and it matches
only an identical-looking expression.
Note that \texttt{match\_dup} should not be used to tell the compiler that a particular
register is being used for two operands (example: \texttt{add} that adds one register to
another; the second register is both an input operand and the output operand).
Use a matching constraint (see Section 16.8.1 [Simple Constraints], page 302)
for those. \texttt{match\_dup} is for the cases where one operand is used in two places
in the template, such as an instruction that computes both a quotient and a
remainder, where the opcode takes two input operands but the RTL template
has to refer to each of those twice; once for the quotient pattern and once for
the remainder pattern.

\texttt{(match\_operator:mm n predicate [operands...])}
This pattern is a kind of placeholder for a variable RTL expression code.
When constructing an insn, it stands for an RTL expression whose expression
code is taken from that of operand \texttt{n}, and whose operands are constructed from
the patterns \texttt{operands}.
When matching an expression, it matches an expression if the function `predicate` returns nonzero on that expression and the patterns `operands` match the operands of the expression.

Suppose that the function `commutative_operator` is defined as follows, to match any expression whose operator is one of the commutative arithmetic operators of RTL and whose mode is `mode`:

```c
int commutative_integer_operator (x, mode)
    rtx x;
    enum machine_mode mode;
{
    enum rtx_code code = GET_CODE (x);
    if (GET_MODE (x) != mode)
        return 0;
    return (GET_RTX_CLASS (code) == RTX_COMM_ARITH
        || code == EQ || code == NE);
}
```

Then the following pattern will match any RTL expression consisting of a commutative operator applied to two general operands:

```plaintext
(match_operator:SI 3 "commutative_operator"
 [(match_operand:SI 1 "general_operand" "g")
  (match_operand:SI 2 "general_operand" "g")])
```

Here the vector `[operands ...]` contains two patterns because the expressions to be matched all contain two operands.

When this pattern does match, the two operands of the commutative operator are recorded as operands 1 and 2 of the insn. (This is done by the two instances of `match_operand`.) Operand 3 of the insn will be the entire commutative expression: use `GET_CODE (operands[3])` to see which commutative operator was used.

The machine mode `m` of `match_operator` works like that of `match_operand`: it is passed as the second argument to the predicate function, and that function is solely responsible for deciding whether the expression to be matched “has” that mode.

When constructing an insn, argument 3 of the gen-function will specify the operation (i.e. the expression code) for the expression to be made. It should be an RTL expression, whose expression code is copied into a new expression whose operands are arguments 1 and 2 of the gen-function. The subexpressions of argument 3 are not used; only its expression code matters.

When `match_operator` is used in a pattern for matching an insn, it usually best if the operand number of the `match_operator` is higher than that of the actual operands of the insn. This improves register allocation because the register allocator often looks at operands 1 and 2 of insns to see if it can do register tying.

There is no way to specify constraints in `match_operator`. The operand of the insn which corresponds to the `match_operator` never has any constraints because it is never reloaded as a whole. However, if parts of its `operands` are matched by `match_operand` patterns, those parts may have constraints of their own.
(match_op_dup: m n [operands...])
Like match_dup, except that it applies to operators instead of operands. When constructing an insn, operand number n will be substituted at this point. But in matching, match_op_dup behaves differently. It assumes that operand number n has already been determined by a match_operator appearing earlier in the recognition template, and it matches only an identical-looking expression.

(match_parallel n predicate [subpat...])
This pattern is a placeholder for an insn that consists of a parallel expression with a variable number of elements. This expression should only appear at the top level of an insn pattern.

When constructing an insn, operand number n will be substituted at this point. When matching an insn, it matches if the body of the insn is a parallel expression with at least as many elements as the vector of subpat expressions in the match_parallel, if each subpat matches the corresponding element of the parallel, and the function predicate returns nonzero on the parallel that is the body of the insn. It is the responsibility of the predicate to validate elements of the parallel beyond those listed in the match_parallel.

A typical use of match_parallel is to match load and store multiple expressions, which can contain a variable number of elements in a parallel. For example,

(define_insn ""
[(match_parallel 0 "load_multiple_operation"
 (set (match_operand: SI 1 "gpc_reg_operand" "=r")
 (match_operand: SI 2 "memory_operand" "m"))
 (use (reg: SI 179))
 (clobber (reg: SI 179)))]
"
"loadm 0,0,%1,%2")

This example comes from ‘a29k.md’. The function load_multiple_operation is defined in ‘a29k.c’ and checks that subsequent elements in the parallel are the same as the set in the pattern, except that they are referencing subsequent registers and memory locations.

An insn that matches this pattern might look like:

(parallel
 [(set (reg: SI 20) (mem: SI (reg: SI 100)))
 (use (reg: SI 179))
 (clobber (reg: SI 179))
 (set (reg: SI 21)
 (mem: SI (plus: SI (reg: SI 100)
 (const_int 4))))
 (set (reg: SI 22)
 (mem: SI (plus: SI (reg: SI 100)
 (const_int 8))))]

(match_par_dup n [subpat...])
Like match_op_dup, but for match_parallel instead of match_operator.
16.5 Output Templates and Operand Substitution

The output template is a string which specifies how to output the assembler code for an instruction pattern. Most of the template is a fixed string which is output literally. The character ‘%’ is used to specify where to substitute an operand; it can also be used to identify places where different variants of the assembler require different syntax.

In the simplest case, a ‘%’ followed by a digit $n$ says to output operand $n$ at that point in the string.

‘%’ followed by a letter and a digit says to output an operand in an alternate fashion. Four letters have standard, built-in meanings described below. The machine description macro PRINT_OPERAND can define additional letters with nonstandard meanings.

‘%c$digit$’ can be used to substitute an operand that is a constant value without the syntax that normally indicates an immediate operand.

‘%n$digit$’ is like ‘%c$digit$’ except that the value of the constant is negated before printing.

‘%a$digit$’ can be used to substitute an operand as if it were a memory reference, with the actual operand treated as the address. This may be useful when outputting a “load address” instruction, because often the assembler syntax for such an instruction requires you to write the operand as if it were a memory reference.

‘%l$digit$’ is used to substitute a label_ref into a jump instruction.

‘%‘ outputs a number which is unique to each instruction in the entire compilation. This is useful for making local labels to be referred to more than once in a single template that generates multiple assembler instructions.

‘%’ followed by a punctuation character specifies a substitution that does not use an operand. Only one case is standard: ‘%%’ outputs a ‘%’ into the assembler code. Other nonstandard cases can be defined in the PRINT_OPERAND macro. You must also define which punctuation characters are valid with the PRINT_OPERAND_PUNCT_VALID_P macro.

The template may generate multiple assembler instructions. Write the text for the instructions, with ‘\;’ between them.

When the RTL contains two operands which are required by constraint to match each other, the output template must refer only to the lower-numbered operand. Matching operands are not always identical, and the rest of the compiler arranges to put the proper RTL expression for printing into the lower-numbered operand.

One use of nonstandard letters or punctuation following ‘%’ is to distinguish between different assembler languages for the same machine; for example, Motorola syntax versus MIT syntax for the 68000. Motorola syntax requires periods in most opcode names, while MIT syntax does not. For example, the opcode ‘movel’ in MIT syntax is ‘move.l’ in Motorola syntax. The same file of patterns is used for both kinds of output syntax, but the character sequence ‘.’ is used in each place where Motorola syntax wants a period. The PRINT_OPERAND macro for Motorola syntax defines the sequence to output a period; the macro for MIT syntax defines it to do nothing.

As a special case, a template consisting of the single character # instructs the compiler to first split the insn, and then output the resulting instructions separately. This helps eliminate redundancy in the output templates. If you have a define_insn that needs to emit multiple assembler instructions, and there is a matching define_split already
defined, then you can simply use `#` as the output template instead of writing an output template that emits the multiple assembler instructions.

If the macro `ASSEMBLER_DIALECT` is defined, you can use construct of the form `{option0|option1|option2}` in the templates. These describe multiple variants of assembler language syntax. See Section 17.21.7 [Instruction Output], page 516.

### 16.6 C Statements for Assembler Output

Often a single fixed template string cannot produce correct and efficient assembler code for all the cases that are recognized by a single instruction pattern. For example, the opcodes may depend on the kinds of operands; or some unfortunate combinations of operands may require extra machine instructions.

If the output control string starts with a `@`, then it is actually a series of templates, each on a separate line. (Blank lines and leading spaces and tabs are ignored.) The templates correspond to the pattern’s constraint alternatives (see Section 16.8.2 [Multi-Alternative], page 306). For example, if a target machine has a two-address add instruction ‘addr’ to add into a register and another ‘addm’ to add a register to memory, you might write this pattern:

```c
(define_insn "addsi3"
 [(set (match_operand:SI 0 "general_operand" "+r,m")
      (plus:SI (match_operand:SI 1 "general_operand" "+0,0")
               (match_operand:SI 2 "general_operand" "+g,r")))]
 ""
 "@
 addr %2,%0
 addm %2,%0")
```

If the output control string starts with a `*`, then it is not an output template but rather a piece of C program that should compute a template. It should execute a `return` statement to return the template-string you want. Most such templates use C string literals, which require doublequote characters to delimit them. To include these doublequote characters in the string, prefix each one with `\`.

If the output control string is written as a brace block instead of a double-quoted string, it is automatically assumed to be C code. In that case, it is not necessary to put in a leading asterisk, or to escape the doublequotes surrounding C string literals.

The operands may be found in the array `operands`, whose C data type is `rtx []`.

It is very common to select different ways of generating assembler code based on whether an immediate operand is within a certain range. Be careful when doing this, because the result of `INTVAL` is an integer on the host machine. If the host machine has more bits in an `int` than the target machine has in the mode in which the constant will be used, then some of the bits you get from `INTVAL` will be superfluous. For proper results, you must carefully disregard the values of those bits.

It is possible to output an assembler instruction and then go on to output or compute more of them, using the subroutine `output_asm_insn`. This receives two arguments: a template-string and a vector of operands. The vector may be `operands`, or it may be another array of `rtx` that you declare locally and initialize yourself.

When an insn pattern has multiple alternatives in its constraints, often the appearance of the assembler code is determined mostly by which alternative was matched. When this
is so, the C code can test the variable `which_alternative`, which is the ordinal number of the alternative that was actually satisfied (0 for the first, 1 for the second alternative, etc.).

For example, suppose there are two opcodes for storing zero, `clrreg` for registers and `clrmem` for memory locations. Here is how a pattern could use `which_alternative` to choose between them:

```c
(define_insn ""
  [(set (match_operand:SI 0 "general_operand" "+r,m")
        (const_int 0))]
   ""
  {
    return (which_alternative == 0
            ? "clrreg %0" : "clrmem %0");
  })
```

The example above, where the assembler code to generate was *solely* determined by the alternative, could also have been specified as follows, having the output control string start with a `'@'`:

```c
(define_insn ""
  [(set (match_operand:SI 0 "general_operand" "+r,m")
        (const_int 0))]
   ""
   "@
    clrreg %0
    clrmem %0")
```

### 16.7 Predicates

A predicate determines whether a `match_operand` or `match_operator` expression matches, and therefore whether the surrounding instruction pattern will be used for that combination of operands. GCC has a number of machine-independent predicates, and you can define machine-specific predicates as needed. By convention, predicates used with `match_operand` have names that end in `'_operand'`, and those used with `match_operator` have names that end in `'_operator'`.

All predicates are Boolean functions (in the mathematical sense) of two arguments: the RTL expression that is being considered at that position in the instruction pattern, and the machine mode that the `match_operand` or `match_operator` specifies. In this section, the first argument is called `op` and the second argument `mode`. Predicates can be called from C as ordinary two-argument functions; this can be useful in output templates or other machine-specific code.

Operand predicates can allow operands that are not actually acceptable to the hardware, as long as the constraints give reload the ability to fix them up (see Section 16.8 [Constraints], page 302). However, GCC will usually generate better code if the predicates specify the requirements of the machine instructions as closely as possible. Reload cannot fix up operands that must be constants (“immediate operands”); you must use a predicate that allows only constants, or else enforce the requirement in the extra condition.

Most predicates handle their `mode` argument in a uniform manner. If `mode` is `VOIDmode` (unspecified), then `op` can have any mode. If `mode` is anything else, then `op` must have the same mode, unless `op` is a `CONST_INT` or integer `CONST_DOUBLE`. These RTL expressions always have `VOIDmode`, so it would be counterproductive to check that their mode matches.
Instead, predicates that accept \texttt{CONST	extunderscore INT} and/or integer \texttt{CONST	extunderscore DOUBLE} check that the value stored in the constant will fit in the requested mode.

Predicates with this behavior are called \textit{normal}. \texttt{genrecog} can optimize the instruction recognizer based on knowledge of how normal predicates treat modes. It can also diagnose certain kinds of common errors in the use of normal predicates; for instance, it is almost always an error to use a normal predicate without specifying a mode.

Predicates that do something different with their \texttt{mode} argument are called \textit{special}. The generic predicates \texttt{address	extunderscore operand} and \texttt{pmode	extunderscore register	extunderscore operand} are special predicates. \texttt{genrecog} does not do any optimizations or diagnosis when special predicates are used.

### 16.7.1 Machine-Independent Predicates

These are the generic predicates available to all back ends. They are defined in ‘\texttt{recog	extunderscore c}’.

The first category of predicates allow only constant, or \textit{immediate}, operands.

- **immediate	extunderscore operand** [Function]
  
  This predicate allows any sort of constant that fits in \texttt{mode}. It is an appropriate choice for instructions that take operands that must be constant.

- **const	extunderscore int	extunderscore operand** [Function]
  
  This predicate allows any \texttt{CONST	extunderscore INT} expression that fits in \texttt{mode}. It is an appropriate choice for an immediate operand that does not allow a symbol or label.

- **const	extunderscore double	extunderscore operand** [Function]
  
  This predicate accepts any \texttt{CONST	extunderscore DOUBLE} expression that has exactly \texttt{mode}. If \texttt{mode} is \texttt{VOIDmode}, it will also accept \texttt{CONST	extunderscore INT}. It is intended for immediate floating point constants.

The second category of predicates allow only some kind of machine register.

- **register	extunderscore operand** [Function]
  
  This predicate allows any \texttt{REG} or \texttt{SUBREG} expression that is valid for \texttt{mode}. It is often suitable for arithmetic instruction operands on a RISC machine.

- **pmode	extunderscore register	extunderscore operand** [Function]
  
  This is a slight variant on \texttt{register	extunderscore operand} which works around a limitation in the machine-description reader.

  \[
  \text{match	extunderscore operand\ n \ "pmode	extunderscore register	extunderscore operand\" constraint}\]

  means exactly what

  \[
  \text{match	extunderscore operand\:P \ n \ "register	extunderscore operand\" constraint}\]

  would mean, if the machine-description reader accepted ‘\texttt{:P}’ mode suffixes. Unfortunately, it cannot, because \texttt{Pmode} is an alias for some other mode, and might vary with machine-specific options. See Section 17.31 [Misc], page 542.

- **scratch	extunderscore operand** [Function]
  
  This predicate allows hard registers and \texttt{SCRATCH} expressions, but not pseudo-registers. It is used internally by \texttt{match	extunderscore scratch}; it should not be used directly.

The third category of predicates allow only some kind of memory reference.
memory_operand
This predicate allows any valid reference to a quantity of mode \textit{mode} in memory, as determined by the weak form of \texttt{GO\_IF\_LEGITIMATE\_ADDRESS} (see Section 17.14 [Addressing Modes], page 469).

address_operand
This predicate is a little unusual; it allows any operand that is a valid expression for the \textit{address} of a quantity of mode \textit{mode}, again determined by the weak form of \texttt{GO\_IF\_LEGITIMATE\_ADDRESS}. To first order, if `(mem:mode (exp))' is acceptable to \textit{memory_operand}, then \textit{exp} is acceptable to \textit{address_operand}. Note that \textit{exp} does not necessarily have the mode \textit{mode}.

indirect_operand
This is a stricter form of \textit{memory_operand} which allows only memory references with a \textit{general_operand} as the address expression. New uses of this predicate are discouraged, because \textit{general_operand} is very permissive, so it’s hard to tell what an \textit{indirect_operand} does or does not allow. If a target has different requirements for memory operands for different instructions, it is better to define target-specific predicates which enforce the hardware’s requirements explicitly.

push_operand
This predicate allows a memory reference suitable for pushing a value onto the stack. This will be a \texttt{MEM} which refers to \texttt{stack\_pointer\_rtx}, with a side-effect in its address expression (see Section 10.16 [Incdec], page 164); which one is determined by the \texttt{STACK\_PUSH\_CODE} macro (see Section 17.10.1 [Frame Layout], page 436).

pop_operand
This predicate allows a memory reference suitable for popping a value off the stack. Again, this will be a \texttt{MEM} referring to \texttt{stack\_pointer\_rtx}, with a side-effect in its address expression. However, this time \texttt{STACK\_POP\_CODE} is expected.

The fourth category of predicates allow some combination of the above operands.

nonmemory_operand
This predicate allows any immediate or register operand valid for \textit{mode}.

nonimmediate_operand
This predicate allows any register or memory operand valid for \textit{mode}.

general_operand
This predicate allows any immediate, register, or memory operand valid for \textit{mode}.

Finally, there are two generic operator predicates.

comparison_operator
This predicate matches any expression which performs an arithmetic comparison in \textit{mode}; that is, \texttt{COMPARISON\_P} is true for the expression code.

ordered_comparison_operator
This predicate matches any expression which performs an arithmetic comparison in \textit{mode} and whose expression code is valid for integer modes; that is, the expression code will be one of \texttt{eq}, \texttt{ne}, \texttt{lt}, \texttt{ltu}, \texttt{le}, \texttt{leu}, \texttt{gt}, \texttt{gtu}, \texttt{ge}, \texttt{geu}. 
16.7.2 Defining Machine-Specific Predicates

Many machines have requirements for their operands that cannot be expressed precisely using the generic predicates. You can define additional predicates using `define_predicate` and `define_special_predicate` expressions. These expressions have three operands:

- The name of the predicate, as it will be referred to in `match_operand` or `match_operator` expressions.
- An RTL expression which evaluates to true if the predicate allows the operand `op`, false if it does not. This expression can only use the following RTL codes:

  **MATCH_OPERAND**
  When written inside a predicate expression, a `MATCH_OPERAND` expression evaluates to true if the predicate it names would allow `op`. The operand number and constraint are ignored. Due to limitations in `genrecog`, you can only refer to generic predicates and predicates that have already been defined.

  **MATCH_CODE**
  This expression evaluates to true if `op` or a specified subexpression of `op` has one of a given list of RTX codes.

  The first operand of this expression is a string constant containing a comma-separated list of RTX code names (in lower case). These are the codes for which the `MATCH_CODE` will be true.

  The second operand is a string constant which indicates what subexpression of `op` to examine. If it is absent or the empty string, `op` itself is examined. Otherwise, the string constant must be a sequence of digits and/or lowercase letters. Each character indicates a subexpression to extract from the current expression; for the first character this is `op`, for the second and subsequent characters it is the result of the previous character. A digit `n` extracts `XEXP(e, n)`; a letter `l` extracts `XVECEXP(e, 0, n)` where `n` is the alphabetic ordinal of `l` (0 for ‘a’, 1 for ‘b’, and so on). The `MATCH_CODE` then examines the RTX code of the subexpression extracted by the complete string. It is not possible to extract components of an `rtvec` that is not at position 0 within its RTX object.

  **MATCH_TEST**
  This expression has one operand, a string constant containing a C expression. The predicate’s arguments, `op` and `mode`, are available with those names in the C expression. The `MATCH_TEST` evaluates to true if the C expression evaluates to a nonzero value. `MATCH_TEST` expressions must not have side effects.

  **AND**
  **IOR**
  **NOT**

  **IF_THEN_ELSE**
  The basic ‘MATCH_’ expressions can be combined using these logical operators, which have the semantics of the C operators ‘&&’, ‘||’, ‘!’’, and ‘? :’
respectively. As in Common Lisp, you may give an \texttt{AND} or \texttt{IOR} expression an arbitrary number of arguments; this has exactly the same effect as writing a chain of two-argument \texttt{AND} or \texttt{IOR} expressions.

- An optional block of C code, which should execute \texttt{return true} if the predicate is found to match and \texttt{return false} if it does not. It must not have any side effects. The predicate arguments, \texttt{op} and \texttt{mode}, are available with those names.

If a code block is present in a predicate definition, then the RTL expression must evaluate to true and the code block must execute \texttt{return true} for the predicate to allow the operand. The RTL expression is evaluated first; do not re-check anything in the code block that was checked in the RTL expression.

The program \texttt{genrecog} scans \texttt{define_predicate} and \texttt{define_special_predicate} expressions to determine which RTX codes are possibly allowed. You should always make this explicit in the RTL predicate expression, using \texttt{MATCH_OPERAND} and \texttt{MATCH_CODE}.

Here is an example of a simple predicate definition, from the IA64 machine description:

\begin{verbatim}
;; True if \texttt{op} is a \texttt{SYMBOL_REF} which refers to the sdata section.
(define_predicate "small_addr_symbolic_operand"
 (and (match_code "symbol_ref")
  (match_test "SYMBOL_REF_SMALL_ADDR_P (op)")))
\end{verbatim}

And here is another, showing the use of the C block.

\begin{verbatim}
;; True if \texttt{op} is a register operand that is (or could be) a GR reg.
(define Predicate "gr_register_operand"
 (match_operand 0 "register_operand")
 { unsigned int regno;
  if (GET_CODE (op) == SUBREG)
   op = SUBREG_REG (op);
  regno = REGNO (op);
  return (regno >= FIRST_PSEUDO_REGISTER || GENERAL_REGNO_P (regno));
 })
\end{verbatim}

Predicates written with \texttt{define_predicate} automatically include a test that \texttt{mode} is \texttt{VOIDmode}, or \texttt{op} has the same mode as \texttt{mode}, or \texttt{op} is a \texttt{CONST\_INT} or \texttt{CONST\_DOUBLE}. They do not check specifically for integer \texttt{CONST\_DOUBLE}, nor do they test that the value of either kind of constant fits in the requested mode. This is because target-specific predicates that take constants usually have to do more stringent value checks anyway. If you need the exact same treatment of \texttt{CONST\_INT} or \texttt{CONST\_DOUBLE} that the generic predicates provide, use a \texttt{MATCH\_OPERAND} subexpression to call \texttt{const_int_operand}, \texttt{const_double_operand}, or \texttt{immediate_operand}.

Predicates written with \texttt{define_special_predicate} do not get any automatic mode checks, and are treated as having special mode handling by \texttt{genrecog}.

The program \texttt{genpreds} is responsible for generating code to test predicates. It also writes a header file containing function declarations for all machine-specific predicates. It is not necessary to declare these predicates in \texttt{cpu-protos.h}.
16.8 Operand Constraints

Each `match_operand` in an instruction pattern can specify constraints for the operands allowed. The constraints allow you to fine-tune matching within the set of operands allowed by the predicate.

Constraints can say whether an operand may be in a register, and which kinds of register; whether the operand can be a memory reference, and which kinds of address; whether the operand may be an immediate constant, and which possible values it may have. Constraints can also require two operands to match. Side-effects aren’t allowed in operands of inline `asm`, unless ‘<’ or ‘>’ constraints are used, because there is no guarantee that the side-effects will happen exactly once in an instruction that can update the addressing register.

16.8.1 Simple Constraints

The simplest kind of constraint is a string full of letters, each of which describes one kind of operand that is permitted. Here are the letters that are allowed:

- `whitespace`
  
  Whitespace characters are ignored and can be inserted at any position except the first. This enables each alternative for different operands to be visually aligned in the machine description even if they have different number of constraints and modifiers.

- `m`
  
  A memory operand is allowed, with any kind of address that the machine supports in general. Note that the letter used for the general memory constraint can be re-defined by a back end using the `TARGET_MEM_CONSTRAINT` macro.

- `o`
  
  A memory operand is allowed, but only if the address is offsettable. This means that adding a small integer (actually, the width in bytes of the operand, as determined by its machine mode) may be added to the address and the result is also a valid memory address.

  For example, an address which is constant is offsettable; so is an address that is the sum of a register and a constant (as long as a slightly larger constant is also within the range of address-offsets supported by the machine); but an autoincrement or autodecrement address is not offsettable. More complicated indirect/indexed addresses may or may not be offsettable depending on the other addressing modes that the machine supports.

  Note that in an output operand which can be matched by another operand, the constraint letter ‘o’ is valid only when accompanied by both ‘<’ (if the target machine has predecrement addressing) and ‘>’ (if the target machine has preincrement addressing).

- `V`
  
  A memory operand that is not offsettable. In other words, anything that would fit the ‘m’ constraint but not the ‘o’ constraint.

- `<`
  
  A memory operand with autodecrement addressing (either predecrement or postdecrement) is allowed. In inline `asm` this constraint is only allowed if the operand is used exactly once in an instruction that can handle the side-effects. Not using an operand with ‘<’ in constraint string in the inline `asm` pattern at all or using it in multiple instructions isn’t valid, because the side-effects
wouldn’t be performed or would be performed more than once. Furthermore, on some targets the operand with ‘<’ in constraint string must be accompanied by special instruction suffixes like `%00 instruction suffix on PowerPC or `%P0 on IA-64.

‘>’ A memory operand with autoincrement addressing (either preincrement or postincrement) is allowed. In inline `asm the same restrictions as for ‘<’ apply.

‘r’ A register operand is allowed provided that it is in a general register.

‘i’ An immediate integer operand (one with constant value) is allowed. This includes symbolic constants whose values will be known only at assembly time or later.

‘n’ An immediate integer operand with a known numeric value is allowed. Many systems cannot support assembly-time constants for operands less than a word wide. Constraints for these operands should use ‘n’ rather than ‘i’.

‘I’, ‘J’, ‘K’, . . . ‘P’ Other letters in the range ‘I’ through ‘P’ may be defined in a machine-dependent fashion to permit immediate integer operands with explicit integer values in specified ranges. For example, on the 68000, ‘I’ is defined to stand for the range of values 1 to 8. This is the range permitted as a shift count in the shift instructions.

‘E’ An immediate floating operand (expression code `const_double`) is allowed, but only if the target floating point format is the same as that of the host machine (on which the compiler is running).

‘F’ An immediate floating operand (expression code `const_double` or `const_vector`) is allowed.

‘G’, ‘H’ ‘G’ and ‘H’ may be defined in a machine-dependent fashion to permit immediate floating operands in particular ranges of values.

‘s’ An immediate integer operand whose value is not an explicit integer is allowed. This might appear strange; if an insn allows a constant operand with a value not known at compile time, it certainly must allow any known value. So why use ‘s’ instead of ‘i’? Sometimes it allows better code to be generated.

For example, on the 68000 in a fullword instruction it is possible to use an immediate operand; but if the immediate value is between −128 and 127, better code results from loading the value into a register and using the register. This is because the load into the register can be done with a `moveq` instruction. We arrange for this to happen by defining the letter ‘K’ to mean “any integer outside the range −128 to 127”, and then specifying ‘ks’ in the operand constraints.

‘g’ Any register, memory or immediate integer operand is allowed, except for registers that are not general registers.

‘X’ Any operand whatsoever is allowed, even if it does not satisfy `general_operand`. This is normally used in the constraint of a `match_scratch` when certain alternatives will not actually require a scratch register.
An operand that matches the specified operand number is allowed. If a digit is used together with letters within the same alternative, the digit should come last.

This number is allowed to be more than a single digit. If multiple digits are encountered consecutively, they are interpreted as a single decimal integer. There is scant chance for ambiguity, since to-date it has never been desirable that ‘10’ be interpreted as matching either operand 1 or operand 0. Should this be desired, one can use multiple alternatives instead.

This is called a matching constraint and what it really means is that the assembler has only a single operand that fills two roles considered separate in the RTL insn. For example, an add insn has two input operands and one output operand in the RTL, but on most CISC machines an add instruction really has only two operands, one of them an input-output operand:

```
addl #35,r12
```

Matching constraints are used in these circumstances. More precisely, the two operands that match must include one input-only operand and one output-only operand. Moreover, the digit must be a smaller number than the number of the operand that uses it in the constraint.

For operands to match in a particular case usually means that they are identical-looking RTL expressions. But in a few special cases specific kinds of dissimilarity are allowed. For example, *x as an input operand will match *x++ as an output operand. For proper results in such cases, the output template should always use the output-operand’s number when printing the operand.

An operand that is a valid memory address is allowed. This is for “load address” and “push address” instructions.

`p` in the constraint must be accompanied by `address_operand` as the predicate in the `match_operand`. This predicate interprets the mode specified in the `match_operand` as the mode of the memory reference for which the address would be valid.

Other letters can be defined in machine-dependent fashion to stand for particular classes of registers or other arbitrary operand types. ‘d’, ‘a’ and ‘f’ are defined on the 68000/68020 to stand for data, address and floating point registers.

In order to have valid assembler code, each operand must satisfy its constraint. But a failure to do so does not prevent the pattern from applying to an insn. Instead, it directs the compiler to modify the code so that the constraint will be satisfied. Usually this is done by copying an operand into a register.

Contrast, therefore, the two instruction patterns that follow:

```lisp
(define_insn ""
 [(set (match_operand:SI 0 "general_operand" ";r")
   (plus:SI (match_dup 0)
     (match_operand:SI 1 "general_operand" ";r")))]
"
```
which has two operands, one of which must appear in two places, and

\[
\text{(define_insn "
    [(set (match_operand:SI 0 "general_operand" ";r")
           (plus:SI (match_operand:SI 1 "general_operand" ";0")
             (match_operand:SI 2 "general_operand" ";r")))
    ""
    "")}
\]

which has three operands, two of which are required by a constraint to be identical. If we are considering an insn of the form

\[
\text{(insn n prev next}
  \text{(set (reg:SI 3)}
    \text{(plus:SI (reg:SI 1) (reg:SI 109))}
  \text{...)}
\]

the first pattern would not apply at all, because this insn does not contain two identical subexpressions in the right place. The pattern would say, “That does not look like an add instruction; try other patterns”. The second pattern would say, “Yes, that’s an add instruction, but there is something wrong with it”. It would direct the reload pass of the compiler to generate additional insns to make the constraint true. The results might look like this:

\[
\text{(insn n2 prev n}
  \text{(set (reg:SI 3) (reg:SI 6))}
  \text{...)}
\]

\[
\text{(insn n n2 next}
  \text{(set (reg:SI 3)}
    \text{(plus:SI (reg:SI 1) (reg:SI 109))}
  \text{...)}
\]

It is up to you to make sure that each operand, in each pattern, has constraints that can handle any RTL expression that could be present for that operand. (When multiple alternatives are in use, each pattern must, for each possible combination of operand expressions, have at least one alternative which can handle that combination of operands.) The constraints don’t need to allow any possible operand—when this is the case, they do not constrain—but they must at least point the way to reloading any possible operand so that it will fit.

- If the constraint accepts whatever operands the predicate permits, there is no problem: reloading is never necessary for this operand.

For example, an operand whose constraints permit everything except registers is safe provided its predicate rejects registers.

An operand whose predicate accepts only constant values is safe provided its constraints include the letter ‘\text{i}’. If any possible constant value is accepted, then nothing less than ‘\text{i}’ will do; if the predicate is more selective, then the constraints may also be more selective.

- Any operand expression can be reloaded by copying it into a register. So if an operand’s constraints allow some kind of register, it is certain to be safe. It need not permit all classes of registers; the compiler knows how to copy a register into another register of the proper class in order to make an instruction valid.
• A nonoffsettable memory reference can be reloaded by copying the address into a register. So if the constraint uses the letter ‘o’, all memory references are taken care of.

• A constant operand can be reloaded by allocating space in memory to hold it as preinitialized data. Then the memory reference can be used in place of the constant. So if the constraint uses the letters ‘o’ or ‘m’, constant operands are not a problem.

• If the constraint permits a constant and a pseudo register used in an insn was not allocated to a hard register and is equivalent to a constant, the register will be replaced with the constant. If the predicate does not permit a constant and the insn is re-recognized for some reason, the compiler will crash. Thus the predicate must always recognize any objects allowed by the constraint.

If the operand’s predicate can recognize registers, but the constraint does not permit them, it can make the compiler crash. When this operand happens to be a register, the reload pass will be stymied, because it does not know how to copy a register temporarily into memory.

If the predicate accepts a unary operator, the constraint applies to the operand. For example, the MIPS processor at ISA level 3 supports an instruction which adds two registers in SIMode to produce a DImode result, but only if the registers are correctly sign extended. This predicate for the input operands accepts a sign_extend of an SIMode register. Write the constraint to indicate the type of register that is required for the operand of the sign_extend.

16.8.2 Multiple Alternative Constraints

Sometimes a single instruction has multiple alternative sets of possible operands. For example, on the 68000, a logical-or instruction can combine register or an immediate value into memory, or it can combine any kind of operand into a register; but it cannot combine one memory location into another.

These constraints are represented as multiple alternatives. An alternative can be described by a series of letters for each operand. The overall constraint for an operand is made from the letters for this operand from the first alternative, a comma, the letters for this operand from the second alternative, a comma, and so on until the last alternative. Here is how it is done for fullword logical-or on the 68000:

```c
(define_insn "iorsi3"
  [(set (match_operand:SI 0 "general_operand" "=m,d")
       (ior:SI (match_operand:SI 1 "general_operand" "%0,0")
               (match_operand:SI 2 "general_operand" "dKs,dmKs")))])
```

The first alternative has ‘m’ (memory) for operand 0, ‘0’ for operand 1 (meaning it must match operand 0), and ‘dKs’ for operand 2. The second alternative has ‘d’ (data register) for operand 0, ‘0’ for operand 1, and ‘dmKs’ for operand 2. The ‘=’ and ‘%’ in the constraints apply to all the alternatives; their meaning is explained in the next section (see Section 16.8.3 [Class Preferences], page 307).

If all the operands fit any one alternative, the instruction is valid. Otherwise, for each alternative, the compiler counts how many instructions must be added to copy the operands so that that alternative applies. The alternative requiring the least copying is chosen. If
two alternatives need the same amount of copying, the one that comes first is chosen. These choices can be altered with the ‘?’ and ‘!’ characters:

? Disparage slightly the alternative that the ‘?’ appears in, as a choice when no alternative applies exactly. The compiler regards this alternative as one unit more costly for each ‘?’ that appears in it.

! Disparage severely the alternative that the ‘!’ appears in. This alternative can still be used if it fits without reloading, but if reloading is needed, some other alternative will be used.

When an insn pattern has multiple alternatives in its constraints, often the appearance of the assembler code is determined mostly by which alternative was matched. When this is so, the C code for writing the assembler code can use the variable `which_alternative`, which is the ordinal number of the alternative that was actually satisfied (0 for the first, 1 for the second alternative, etc.). See Section 16.6 [Output Statement], page 296.

### 16.8.3 Register Class Preferences

The operand constraints have another function: they enable the compiler to decide which kind of hardware register a pseudo register is best allocated to. The compiler examines the constraints that apply to the insns that use the pseudo register, looking for the machine-dependent letters such as ‘d’ and ‘a’ that specify classes of registers. The pseudo register is put in whichever class gets the most “votes”. The constraint letters ‘g’ and ‘r’ also vote: they vote in favor of a general register. The machine description says which registers are considered general.

Of course, on some machines all registers are equivalent, and no register classes are defined. Then none of this complexity is relevant.

### 16.8.4 Constraint Modifier Characters

Here are constraint modifier characters.

‘=' Means that this operand is write-only for this instruction: the previous value is discarded and replaced by output data.

‘+’ Means that this operand is both read and written by the instruction.

When the compiler fixes up the operands to satisfy the constraints, it needs to know which operands are inputs to the instruction and which are outputs from it. ‘=’ identifies an output; ‘+’ identifies an operand that is both input and output; all other operands are assumed to be input only.

If you specify ‘=’ or ‘+’ in a constraint, you put it in the first character of the constraint string.

‘&’ Means (in a particular alternative) that this operand is an earlyclobber operand, which is modified before the instruction is finished using the input operands. Therefore, this operand may not lie in a register that is used as an input operand or as part of any memory address.

‘&’ applies only to the alternative in which it is written. In constraints with multiple alternatives, sometimes one alternative requires ‘&’ while others do not. See, for example, the ‘movdf’ insn of the 68000.
An input operand can be tied to an earlyclobber operand if its only use as an input occurs before the early result is written. Adding alternatives of this form often allows GCC to produce better code when only some of the inputs can be affected by the earlyclobber. See, for example, the `mulsi3` insn of the ARM. ‘&’ does not obviate the need to write ‘=’.

‘%’

Declares the instruction to be commutative for this operand and the following operand. This means that the compiler may interchanging the two operands if that is the cheapest way to make all operands fit the constraints. This is often used in patterns for addition instructions that really have only two operands: the result must go in one of the arguments. Here for example, is how the 68000 halfword-add instruction is defined:

```
(define_insn "addhi3"
  [(set (match_operand:HI 0 "general_operand" 
                     "="m,r")
         (plus:HI (match_operand:HI 1 "general_operand" 
                          
                     "%0,0")
         (match_operand:HI 2 "general_operand" 
                          "di,g")))]

GCC can only handle one commutative pair in an asm; if you use more, the compiler may fail. Note that you need not use the modifier if the two alternatives are strictly identical; this would only waste time in the reload pass. The modifier is not operational after register allocation, so the result of `define_peephole2` and `define_splits` performed after reload cannot rely on ‘%’ to make the intended insn match.

‘#’

Says that all following characters, up to the next comma, are to be ignored as a constraint. They are significant only for choosing register preferences.

‘*’

Says that the following character should be ignored when choosing register preferences. ‘*’ has no effect on the meaning of the constraint as a constraint, and no effect on reloading.

Here is an example: the 68000 has an instruction to sign-extend a halfword in a data register, and can also sign-extend a value by copying it into an address register. While either kind of register is acceptable, the constraints on an address-register destination are less strict, so it is best if register allocation makes an address register its goal. Therefore, ‘*’ is used so that the ‘d’ constraint letter (for data register) is ignored when computing register preferences.

```
(define_insn "extendhisi2"
  [(set (match_operand:SI 0 "general_operand"  
                     "=*d,a")
         (sign_extend:SI 
                     (match_operand:HI 1 "general_operand"  
                                     "0,g")))]

16.8.5 Constraints for Particular Machines

Whenever possible, you should use the general-purpose constraint letters in *asm* arguments, since they will convey meaning more readily to people reading your code. Failing that, use the constraint letters that usually have very similar meanings across architectures. The most commonly used constraints are ‘m’ and ‘r’ (for memory and general-purpose registers respectively; see Section 16.8.1 [Simple Constraints], page 302), and ‘I’, usually the letter indicating the most common immediate-constant format.
Each architecture defines additional constraints. These constraints are used by the compiler itself for instruction generation, as well as for `asm` statements; therefore, some of the constraints are not particularly useful for `asm`. Here is a summary of some of the machine-dependent constraints available on some particular machines; it includes both constraints that are useful for `asm` and constraints that aren’t. The compiler source file mentioned in the table heading for each architecture is the definitive reference for the meanings of that architecture’s constraints.

**ARM family—`config/arm/arm.h`**

- **f** Floating-point register
- **w** VFP floating-point register
- **F** One of the floating-point constants 0.0, 0.5, 1.0, 2.0, 3.0, 4.0, 5.0 or 10.0
- **G** Floating-point constant that would satisfy the constraint ‘F’ if it were negated
- **I** Integer that is valid as an immediate operand in a data processing instruction. That is, an integer in the range 0 to 255 rotated by a multiple of 2
- **J** Integer in the range −4095 to 4095
- **K** Integer that satisfies constraint ‘I’ when inverted (ones complement)
- **L** Integer that satisfies constraint ‘I’ when negated (twos complement)
- **M** Integer in the range 0 to 32
- **Q** A memory reference where the exact address is in a single register (“m” is preferable for `asm` statements)
- **R** An item in the constant pool
- **S** A symbol in the text segment of the current file
- **Uv** A memory reference suitable for VFP load/store insns (reg+constant offset)
- **Uy** A memory reference suitable for iWMMXt load/store instructions.
- **Uq** A memory reference suitable for the ARMv4 ldrsb instruction.

**AVR family—`config/avr/constraints.md`**

- **l** Registers from r0 to r15
- **a** Registers from r16 to r23
- **d** Registers from r16 to r31
- **w** Registers from r24 to r31. These registers can be used in ‘adiw’ command
- **e** Pointer register (r26–r31)
b Base pointer register (r28–r31)
q Stack pointer register (SPH:SPL)
t Temporary register r0
x Register pair X (r27:r26)
y Register pair Y (r29:r28)
z Register pair Z (r31:r30)
I Constant greater than −1, less than 64
J Constant greater than −64, less than 1
K Constant integer 2
L Constant integer 0
M Constant that fits in 8 bits
N Constant integer −1
O Constant integer 8, 16, or 24
P Constant integer 1
G A floating point constant 0.0
R Integer constant in the range −6 . . . 5.
Q A memory address based on Y or Z pointer with displacement.

**CRX Architecture—`config/crx/crx.h`**

b Registers from r0 to r14 (registers without stack pointer)
l Register r16 (64-bit accumulator lo register)
h Register r17 (64-bit accumulator hi register)
k Register pair r16-r17. (64-bit accumulator lo-hi pair)
I Constant that fits in 3 bits
J Constant that fits in 4 bits
K Constant that fits in 5 bits
L Constant that is one of −1, 4, −4, 7, 8, 12, 16, 20, 32, 48
G Floating point constant that is legal for store immediate

**Hewlett-Packard PA-RISC—`config/pa/pa.h`**
a General register 1
f Floating point register
q Shift amount register
x Floating point register (deprecated)
y Upper floating point register (32-bit), floating point register (64-bit)
Z  Any register
I  Signed 11-bit integer constant
J  Signed 14-bit integer constant
K  Integer constant that can be deposited with a \texttt{zdepi} instruction
L  Signed 5-bit integer constant
M  Integer constant 0
N  Integer constant that can be loaded with a \texttt{ldil} instruction
O  Integer constant whose value plus one is a power of 2
P  Integer constant that can be used for \texttt{and} operations in \texttt{depi} and \texttt{extru} instructions
S  Integer constant 31
U  Integer constant 63
G  Floating-point constant 0.0
A  A \texttt{lo_sum} data-linkage-table memory operand
Q  A memory operand that can be used as the destination operand of an integer store instruction
R  A scaled or unscaled indexed memory operand
T  A memory operand for floating-point loads and stores
W  A register indirect memory operand

\textit{picoChip family—'picochip.h'}

k  Stack register.
f  Pointer register. A register which can be used to access memory without supplying an offset. Any other register can be used to access memory, but will need a constant offset. In the case of the offset being zero, it is more efficient to use a pointer register, since this reduces code size.
t  A twin register. A register which may be paired with an adjacent register to create a 32-bit register.
a  Any absolute memory address (e.g., symbolic constant, symbolic constant + offset).
I  4-bit signed integer.
J  4-bit unsigned integer.
K  8-bit signed integer.
M  Any constant whose absolute value is no greater than 4-bits.
N  10-bit signed integer
16-bit signed integer.

**PowerPC and IBM RS6000**—`config/rs6000/rs6000.h`

- **b** Address base register
- **d** Floating point register (containing 64-bit value)
- **f** Floating point register (containing 32-bit value)
- **v** Altivec vector register
- **wd** VSX vector register to hold vector double data
- **wf** VSX vector register to hold vector float data
- **ws** VSX vector register to hold scalar float data
- **wa** Any VSX register
- **h** ‘MQ’, ‘CTR’, or ‘LINK’ register
- **q** ‘MQ’ register
- **c** ‘CTR’ register
- **l** ‘LINK’ register
- **x** ‘CR’ register (condition register) number 0
- **y** ‘CR’ register (condition register)
- **z** ‘XER[CA]’ carry bit (part of the XER register)
- **I** Signed 16-bit constant
- **J** Unsigned 16-bit constant shifted left 16 bits (use ‘L’ instead for SImode constants)
- **K** Unsigned 16-bit constant
- **L** Signed 16-bit constant shifted left 16 bits
- **M** Constant larger than 31
- **N** Exact power of 2
- **O** Zero
- **P** Constant whose negation is a signed 16-bit constant
- **G** Floating point constant that can be loaded into a register with one instruction per word
- **H** Integer/Floating point constant that can be loaded into a register using three instructions
- **m** Memory operand. Normally, m does not allow addresses that update the base register. If ‘<’ or ‘>’ constraint is also used, they are allowed and therefore on PowerPC targets in that case it is only safe to use ‘m<’ in an asm statement if that asm statement accesses the operand exactly once. The asm statement must also use ‘%U<opno>’ as a placeholder for the “update” flag in the corresponding load or store instruction. For example:
asm ("st%U0 %1,%0" : "%m<>" (mem) : "r" (val));
is correct but:
asm ("st %1,%0" : "%m<>" (mem) : "r" (val));
is not.

es A “stable” memory operand; that is, one which does not include any automodification of the base register. This used to be useful when 'm' allowed automodification of the base register, but as those are now only allowed when '<' or '>' is used, 'es' is basically the same as 'm' without '<' and '>'.

Q Memory operand that is an offset from a register (it is usually better to use 'm' or 'es' in asm statements)

Z Memory operand that is an indexed or indirect from a register (it is usually better to use 'm' or 'es' in asm statements)

R AIX TOC entry

a Address operand that is an indexed or indirect from a register ('p' is preferable for asm statements)

S Constant suitable as a 64-bit mask operand

T Constant suitable as a 32-bit mask operand

U System V Release 4 small data area reference

t AND masks that can be performed by two rldic{l, r} instructions

W Vector constant that does not require memory

j Vector constant that is all zeros.

Intel 386—`config/i386/constraints.md`

R Legacy register—the eight integer registers available on all i386 processors (a, b, c, d, si, di, bp, sp).

q Any register accessible as r1. In 32-bit mode, a, b, c, and d; in 64-bit mode, any integer register.

Q Any register accessible as rh: a, b, c, and d.

l Any register that can be used as the index in a base+index memory access: that is, any general register except the stack pointer.

a The a register.

b The b register.

c The c register.

d The d register.

S The si register.

D The di register.
The \texttt{a} and \texttt{d} registers. This class is used for instructions that return double word results in the \texttt{ax:dx} register pair. Single word values will be allocated either in \texttt{ax} or \texttt{dx}. For example on \texttt{i386} the following implements \texttt{rdtsc}:

```c
unsigned long long rdtsc (void)
{
    unsigned long long tick;
    __asm__ __volatile__("rdtsc":"=A"(tick));
    return tick;
}
```

This is not correct on \texttt{x86_64} as it would allocate \texttt{tick} in either \texttt{ax} or \texttt{dx}. You have to use the following variant instead:

```c
unsigned long long rdtsc (void)
{
    unsigned int tickl, tickh;
    __asm__ __volatile__("rdtsc":"=a"(tickl),"=d"(tickh));
    return ((unsigned long long)tickh << 32)|tickl;
}
```

\texttt{f} Any 80387 floating-point (stack) register.

\texttt{t} Top of 80387 floating-point stack (%st(0)).

\texttt{u} Second from top of 80387 floating-point stack (%st(1)).

\texttt{y} Any MMX register.

\texttt{x} Any SSE register.

\texttt{Yz} First SSE register (%xmm0).

\texttt{Y2} Any SSE register, when SSE2 is enabled.

\texttt{Yi} Any SSE register, when SSE2 and inter-unit moves are enabled.

\texttt{Ym} Any MMX register, when inter-unit moves are enabled.

\texttt{I} Integer constant in the range 0 \ldots 31, for 32-bit shifts.

\texttt{J} Integer constant in the range 0 \ldots 63, for 64-bit shifts.

\texttt{K} Signed 8-bit integer constant.

\texttt{L} 0xFF or 0xFFFF, for andsi as a zero-extending move.

\texttt{M} 0, 1, 2, or 3 (shifts for the \texttt{lea} instruction).

\texttt{N} Unsigned 8-bit integer constant (for \texttt{in} and \texttt{out} instructions).

\texttt{O} Integer constant in the range 0 \ldots 127, for 128-bit shifts.

\texttt{G} Standard 80387 floating point constant.

\texttt{C} Standard SSE floating point constant.

\texttt{e} 32-bit signed integer constant, or a symbolic reference known to fit that range (for immediate operands in sign-extending x86-64 instructions).
32-bit unsigned integer constant, or a symbolic reference known to fit that range (for immediate operands in zero-extending x86-64 instructions).

**Intel IA-64—`config/ia64/ia64.h`**

<table>
<thead>
<tr>
<th>Letter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>General register r0 to r3 for addl instruction</td>
</tr>
<tr>
<td>b</td>
<td>Branch register</td>
</tr>
<tr>
<td>c</td>
<td>Predicate register (‘c’ as in “conditional”)</td>
</tr>
<tr>
<td>d</td>
<td>Application register residing in M-unit</td>
</tr>
<tr>
<td>e</td>
<td>Application register residing in I-unit</td>
</tr>
<tr>
<td>f</td>
<td>Floating-point register</td>
</tr>
<tr>
<td>m</td>
<td>Memory operand. If used together with ‘&lt;’ or ‘&gt;’, the operand can have postincrement and postdecrement which require printing with ‘%Pn’ on IA-64.</td>
</tr>
<tr>
<td>g</td>
<td>Floating-point constant 0.0 or 1.0</td>
</tr>
<tr>
<td>i</td>
<td>14-bit signed integer constant</td>
</tr>
<tr>
<td>j</td>
<td>22-bit signed integer constant</td>
</tr>
<tr>
<td>k</td>
<td>8-bit signed integer constant for logical instructions</td>
</tr>
<tr>
<td>l</td>
<td>8-bit adjusted signed integer constant for compare pseudo-ops</td>
</tr>
<tr>
<td>m</td>
<td>6-bit unsigned integer constant for shift counts</td>
</tr>
<tr>
<td>n</td>
<td>9-bit signed integer constant for load and store postincrements</td>
</tr>
<tr>
<td>o</td>
<td>The constant zero</td>
</tr>
<tr>
<td>p</td>
<td>0 or −1 for dep instruction</td>
</tr>
<tr>
<td>q</td>
<td>Non-volatile memory for floating-point loads and stores</td>
</tr>
<tr>
<td>r</td>
<td>Integer constant in the range 1 to 4 for shladd instruction</td>
</tr>
<tr>
<td>s</td>
<td>Memory operand except postincrement and postdecrement. This is now roughly the same as ‘m’ when not used together with ‘&lt;’ or ‘&gt;’.</td>
</tr>
</tbody>
</table>

**FRV—`config/frv/frv.h`**

<table>
<thead>
<tr>
<th>Letter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Register in the class ACC_REGS (acc0 to acc7).</td>
</tr>
<tr>
<td>b</td>
<td>Register in the class EVEN_ACC_REGS (acc0 to acc7).</td>
</tr>
<tr>
<td>c</td>
<td>Register in the class CC_REGS (fcc0 to fcc3 and icc0 to icc3).</td>
</tr>
<tr>
<td>d</td>
<td>Register in the class GPR_REGS (gr0 to gr63).</td>
</tr>
<tr>
<td>e</td>
<td>Register in the class EVEN_REGS (gr0 to gr63). Odd registers are excluded not in the class but through the use of a machine mode larger than 4 bytes.</td>
</tr>
</tbody>
</table>
Register in the class **FPR_REGS** (fr0 to fr63).

Register in the class **FEVEN_REGS** (fr0 to fr63). Odd registers are excluded not in the class but through the use of a machine mode larger than 4 bytes.

Register in the class **LR_REG** (the lr register).

Register in the class **QUAD_REGS** (gr2 to gr63). Register numbers not divisible by 4 are excluded not in the class but through the use of a machine mode larger than 8 bytes.

Register in the class **ICC_REGS** (icc0 to icc3).

Register in the class **FCC_REGS** (fcc0 to fcc3).

Register in the class **ICR_REGS** (cc4 to cc7).

Register in the class **FCR_REGS** (cc0 to cc3).

Register in the class **QUAD_FPR_REGS** (fr0 to fr63). Register numbers not divisible by 4 are excluded not in the class but through the use of a machine mode larger than 8 bytes.

Register in the class **SPR_REGS** (lcr and lr).

Register in the class **QUAD_ACC_REGS** (acc0 to acc7).

Register in the class **ACCG_REGS** (accg0 to accg7).

Register in the class **CR_REGS** (cc0 to cc7).

Floating point constant zero

6-bit signed integer constant

10-bit signed integer constant

16-bit signed integer constant

16-bit unsigned integer constant

12-bit signed integer constant that is negative—i.e. in the range of \(-2048\) to \(-1\)

Constant zero

12-bit signed integer constant that is greater than zero—i.e. in the range of 1 to 2047.

*Blackfin family—‘config/bfin/constraints.md’*

A single register. If \(n\) is in the range 0 to 7, the corresponding D register. If it is A, then the register P0.

Even-numbered D register
Odd-numbered D register
Accumulator register.
Even-numbered accumulator register.
Odd-numbered accumulator register.
I register
B register
M register
Registers used for circular buffering, i.e. I, B, or L registers.
The CC register.
LT0 or LT1.
LC0 or LC1.
LB0 or LB1.
Any D, P, B, M, I or L register.
Additional registers typically used only in prologues and epilogues: RETS, RETN, RETI, RETX, RETE, ASTAT, SEQSTAT and USP.
Any register except accumulators or CC.
Signed 16 bit integer (in the range −32768 to 32767)
Unsigned 16 bit integer (in the range 0 to 65535)
Signed 7 bit integer (in the range −64 to 63)
Unsigned 7 bit integer (in the range 0 to 127)
Unsigned 5 bit integer (in the range 0 to 31)
Signed 4 bit integer (in the range −8 to 7)
Signed 3 bit integer (in the range −3 to 4)
Unsigned 3 bit integer (in the range 0 to 7)
Constant n, where n is a single-digit constant in the range 0 to 4.
An integer equal to one of the MACFLAG_XXX constants that is suitable for use with either accumulator.
An integer equal to one of the MACFLAG_XXX constants that is suitable for use only with accumulator A1.
Constant 255.
Constant 65535.
An integer constant with exactly a single bit set.
An integer constant with all bits set except exactly one.
Q       Any SYMBOL_REF.

M32C—‘config/m32c/m32c.c’
Rsp  Rfb  Rsb  ‘$sp’, ‘$fb’, ‘$sb’.
Rcr  Any control register, when they’re 16 bits wide (nothing if control
     registers are 24 bits wide)
Rcl  Any control register, when they’re 24 bits wide.
R0w  R1w  R2w  $r0, $r1, $r2, $r3.
R3w  $r0 or $r2, or $r2$r0 for 32 bit values.
R02  $r0 or $r2, or $r2$r0 for 32 bit values.
R13  $r1 or $r3, or $r3$r1 for 32 bit values.
Rdi  A register that can hold a 64 bit value.
Rhl  $r0 or $r1 (registers with addressable high/low bytes)
R23  $r2 or $r3
Raa  Address registers
Raw  Address registers when they’re 16 bits wide.
Ral  Address registers when they’re 24 bits wide.
Rqi  Registers that can hold QI values.
Rad  Registers that can be used with displacements ($a0, $a1, $sb).
Rsi  Registers that can hold 32 bit values.
Rhi  Registers that can hold 16 bit values.
Rhc  Registers that can hold 16 bit values, including all control registers.
Rra  $r0 through R1, plus $a0 and $a1.
Rfl  The flags register.
Rmm  The memory-based pseudo-registers $mem0 through $mem15.
Rpi  Registers that can hold pointers (16 bit registers for r8c, m16c; 24
     bit registers for m32cm, m32c).
Rpa  Matches multiple registers in a PARALLEL to form a larger regis-
     ter. Used to match function return values.
Is3  −8 . . . 7
IS1  −128 . . . 127
IS2  −32768 . . . 32767
IU2 \( 0 \ldots 65535 \)
In4 \( -8 \ldots -1 \) or \( 1 \ldots 8 \)
In5 \( -16 \ldots -1 \) or \( 1 \ldots 16 \)
In6 \( -32 \ldots -1 \) or \( 1 \ldots 32 \)
IM2 \( -65536 \ldots -1 \)
I1b An 8 bit value with exactly one bit set.
I1w A 16 bit value with exactly one bit set.
Sd The common src/dest memory addressing modes.
Sa Memory addressed using $a0 or $a1.
Si Memory addressed with immediate addresses.
Ss Memory addressed using the stack pointer ($sp$).
Sf Memory addressed using the frame base register ($fb$).
Ss Memory addressed using the small base register ($sb$).
S1 The $r1h$

\textit{MeP—`config/mep/constraints.md’}

a The $sp$ register.
b The $tp$ register.
c Any control register.
d Either the $hi$ or the $lo$ register.
em Coprocessor registers that can be directly loaded ($c0$-$c15$).
ex Coprocessor registers that can be moved to each other.
er Coprocessor registers that can be moved to core registers.
h The $hi$ register.
j The $rpc$ register.
l The $lo$ register.
t Registers which can be used in $tp$-relative addressing.
v The $gp$ register.
x The coprocessor registers.
y The coprocessor control registers.
z The $0$ register.
A User-defined register set A.
B User-defined register set B.
C User-defined register set C.
User-defined register set D.

Offsets for $gp-rel addressing.

Constants that can be used directly with boolean insns.

Constants that can be moved directly to registers.

Small constants that can be added to registers.

Long shift counts.

Small constants that can be compared to registers.

Constants that can be loaded into the top half of registers.

Signed 8-bit immediates.

Symbols encoded for $tp-rel or $gp-rel addressing.

Non-constant addresses for loading/saving coprocessor registers.

The top half of a symbol’s value.

A register indirect address without offset.

Symbolic references to the control bus.

A general register (r0 to r31).

A status register ($rmsr, $fcc1 to $fcc7).

An address register. This is equivalent to r unless generating MIPS16 code.

A floating-point register (if available).

Formerly the hi register. This constraint is no longer supported.

The lo register. Use this register to store values that are no bigger than a word.

The concatenated hi and lo registers. Use this register to store doubleword values.

A register suitable for use in an indirect jump. This will always be $25 for -mabicalls.

Register $3. Do not use this constraint in new code; it is retained only for compatibility with glibc.

Equivalent to r; retained for backwards compatibility.

A floating-point condition code register.

A signed 16-bit constant (for arithmetic instructions).

Integer zero.

An unsigned 16-bit constant (for logic instructions).
L  A signed 32-bit constant in which the lower 16 bits are zero. Such constants can be loaded using lui.
M  A constant that cannot be loaded using lui, addiu or ori.
N  A constant in the range −65535 to −1 (inclusive).
O  A signed 15-bit constant.
P  A constant in the range 1 to 65535 (inclusive).
G  Floating-point zero.
R  An address that can be used in a non-macro load or store.

Motorola 680x0—‘config/m68k/constraints.md’
a Address register
d Data register
f 68881 floating-point register, if available
I Integer in the range 1 to 8
J 16-bit signed number
K Signed number whose magnitude is greater than 0x80
L Integer in the range −8 to −1
M Signed number whose magnitude is greater than 0x100
N Range 24 to 31, rotatert:SI 8 to 1 expressed as rotate
O 16 (for rotate using swap)
P Range 8 to 15, rotatert:HI 8 to 1 expressed as rotate
R Numbers that mov3q can handle
G Floating point constant that is not a 68881 constant
S Operands that satisfy ‘m’ when -mperel is in effect
T Operands that satisfy ‘s’ when -mperel is not in effect
Q Address register indirect addressing mode
U Register offset addressing
W const_call_operand
Cs symbol_ref or const
Ci const_int
C0 const_int 0
Cj Range of signed numbers that don’t fit in 16 bits
Cmvq Integers valid for mvq
Capsw Integers valid for a moveq followed by a swap
Cmvz   Integers valid for mvz
Cmvs   Integers valid for mvs
Ap     push_operand
Ac     Non-register operands allowed in clr

Motorola 68HC11 & 68HC12 families—‘config/m68hc11/m68hc11.h’
a     Register ‘a’
b     Register ‘b’
d     Register ‘d’
q     An 8-bit register
t     Temporary soft register _.tmp
u     A soft register _.d1 to _.d31
w     Stack pointer register
x     Register ‘x’
y     Register ‘y’
z     Pseudo register ‘z’ (replaced by ‘x’ or ‘y’ at the end)
A     An address register: x, y or z
B     An address register: x or y
D     Register pair (x:d) to form a 32-bit value
L     Constants in the range −65536 to 65535
M     Constants whose 16-bit low part is zero
N     Constant integer 1 or −1
O     Constant integer 16
P     Constants in the range −8 to 2

Moxie—‘config/moxie/constraints.md’
A     An absolute address
B     An offset address
W     A register indirect memory operand
I     A constant in the range of 0 to 255.
N     A constant in the range of 0 to −255.

PDP-11—‘config/pdp11/constraints.md’
a     Floating point registers AC0 through AC3. These can be loaded from/to memory with a single instruction.
d     Odd numbered general registers (R1, R3, R5). These are used for 16-bit multiply operations.
Any of the floating point registers (AC0 through AC5).

Floating point constant 0.

An integer constant that fits in 16 bits.

An integer constant whose low order 16 bits are zero.

An integer constant that does not meet the constraints for codes ‘I’ or ‘J’.

The integer constant 1.

The integer constant –1.

The integer constant 0.

Integer constants –4 through –1 and 1 through 4; shifts by these amounts are handled as multiple single-bit shifts rather than a single variable-length shift.

A memory reference which requires an additional word (address or offset) after the opcode.

A memory reference that is encoded within the opcode.

An address which does not involve register indirect addressing or pre/post increment/decrement addressing.

A symbol reference.

A constant in the range –256 to 255, inclusive.

A constant in the range –128 to 127, inclusive.

A constant in the range –32768 to 32767, inclusive.

A constant in the range –8388608 to 8388607, inclusive.

A constant in the range 0 to 15, inclusive.

Floating-point register on the SPARC-V8 architecture and lower floating-point register on the SPARC-V9 architecture.

Floating-point register. It is equivalent to ‘f’ on the SPARC-V8 architecture and contains both lower and upper floating-point registers on the SPARC-V9 architecture.

Floating-point condition code register.

Lower floating-point register. It is only valid on the SPARC-V9 architecture when the Visual Instruction Set is available.

Floating-point register. It is only valid on the SPARC-V9 architecture when the Visual Instruction Set is available.

64-bit global or out register for the SPARC-V8+ architecture.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>A vector constant</td>
</tr>
<tr>
<td>I</td>
<td>Signed 13-bit constant</td>
</tr>
<tr>
<td>J</td>
<td>Zero</td>
</tr>
<tr>
<td>K</td>
<td>32-bit constant with the low 12 bits clear (a constant that can be loaded with the <code>sethi</code> instruction)</td>
</tr>
<tr>
<td>L</td>
<td>A constant in the range supported by <code>movcc</code> instructions</td>
</tr>
<tr>
<td>M</td>
<td>A constant in the range supported by <code>movrcc</code> instructions</td>
</tr>
<tr>
<td>N</td>
<td>Same as 'K', except that it verifies that bits that are not in the lower 32-bit range are all zero. Must be used instead of 'K' for modes wider than <code>SImode</code></td>
</tr>
<tr>
<td>O</td>
<td>The constant 4096</td>
</tr>
<tr>
<td>G</td>
<td>Floating-point zero</td>
</tr>
<tr>
<td>H</td>
<td>Signed 13-bit constant, sign-extended to 32 or 64 bits</td>
</tr>
<tr>
<td>Q</td>
<td>Floating-point constant whose integral representation can be moved into an integer register using a single <code>sethi</code> instruction</td>
</tr>
<tr>
<td>R</td>
<td>Floating-point constant whose integral representation can be moved into an integer register using a single <code>mov</code> instruction</td>
</tr>
<tr>
<td>S</td>
<td>Floating-point constant whose integral representation can be moved into an integer register using a high/lo_sum instruction sequence</td>
</tr>
<tr>
<td>T</td>
<td>Memory address aligned to an 8-byte boundary</td>
</tr>
<tr>
<td>U</td>
<td>Even register</td>
</tr>
<tr>
<td>W</td>
<td>Memory address for 'e' constraint registers</td>
</tr>
<tr>
<td>Y</td>
<td>Vector zero</td>
</tr>
</tbody>
</table>

**SPU**—`config/spu/spu.h`

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>An immediate which can be loaded with the il ila ilh ilhu instructions. <code>const_int</code> is treated as a 64 bit value.</td>
</tr>
<tr>
<td>c</td>
<td>An immediate for and/xor/or instructions. <code>const_int</code> is treated as a 64 bit value.</td>
</tr>
<tr>
<td>d</td>
<td>An immediate for the <code>iohl</code> instruction. <code>const_int</code> is treated as a 64 bit value.</td>
</tr>
<tr>
<td>f</td>
<td>An immediate which can be loaded with <code>fsmbi</code>.</td>
</tr>
<tr>
<td>A</td>
<td>An immediate which can be loaded with the il ila ilh ilhu instructions. <code>const_int</code> is treated as a 32 bit value.</td>
</tr>
<tr>
<td>B</td>
<td>An immediate for most arithmetic instructions. <code>const_int</code> is treated as a 32 bit value.</td>
</tr>
<tr>
<td>C</td>
<td>An immediate for and/xor/or instructions. <code>const_int</code> is treated as a 32 bit value.</td>
</tr>
</tbody>
</table>
D  An immediate for the `iohl` instruction. `const_int` is treated as a 32 bit value.
I  A constant in the range $[-64, 63]$ for shift/rotate instructions.
J  An unsigned 7-bit constant for conversion/nop/channel instructions.
K  A signed 10-bit constant for most arithmetic instructions.
M  A signed 16 bit immediate for `stop`.
N  An unsigned 16-bit constant for `iohl` and `fsmbi`.
O  An unsigned 7-bit constant whose 3 least significant bits are 0.
P  An unsigned 3-bit constant for 16-byte rotates and shifts
R  Call operand, reg, for indirect calls
S  Call operand, symbol, for relative calls.
T  Call operand, `const_int`, for absolute calls.
U  An immediate which can be loaded with the `il ila ilh ilhu` instructions. `const_int` is sign extended to 128 bit.
W  An immediate for shift and rotate instructions. `const_int` is treated as a 32 bit value.
Y  An immediate for `and/xor/or` instructions. `const_int` is sign extended as a 128 bit.
Z  An immediate for the `iohl` instruction. `const_int` is sign extended to 128 bit.

*S/390 and zSeries—`config/s390/s390.h`*

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Address register (general purpose register except r0)</td>
</tr>
<tr>
<td>c</td>
<td>Condition code register</td>
</tr>
<tr>
<td>d</td>
<td>Data register (arbitrary general purpose register)</td>
</tr>
<tr>
<td>f</td>
<td>Floating-point register</td>
</tr>
<tr>
<td>I</td>
<td>Unsigned 8-bit constant (0–255)</td>
</tr>
<tr>
<td>J</td>
<td>Unsigned 12-bit constant (0–4095)</td>
</tr>
<tr>
<td>K</td>
<td>Signed 16-bit constant ($-32768$–32767)</td>
</tr>
<tr>
<td>L</td>
<td>Value appropriate as displacement.</td>
</tr>
<tr>
<td></td>
<td>(0.4095)</td>
</tr>
<tr>
<td></td>
<td>for short displacement</td>
</tr>
<tr>
<td></td>
<td>($-524288$.524287)</td>
</tr>
<tr>
<td></td>
<td>for long displacement</td>
</tr>
<tr>
<td>M</td>
<td>Constant integer with a value of <code>0x7fffffff</code></td>
</tr>
</tbody>
</table>
Multiple letter constraint followed by 4 parameter letters.

- **N**: number of the part counting from most to least significant
- **H, Q**: mode of the part
- **D, S, H**: mode of the containing operand
- **0, F**: value of the other parts (F—all bits set)

The constraint matches if the specified part of a constant has a value different from its other parts.

- **Q**: Memory reference without index register and with short displacement.
- **R**: Memory reference with index register and short displacement.
- **S**: Memory reference without index register but with long displacement.
- **T**: Memory reference with index register and long displacement.
- **U**: Pointer with short displacement.
- **W**: Pointer with long displacement.
- **Y**: Shift count operand.

**Score family**—`config/score/score.h`

- **d**: Registers from r0 to r32.
- **e**: Registers from r0 to r16.
- **t**: r8—r11 or r22—r27 registers.
- **h**: hi register.
- **l**: lo register.
- **x**: hi + lo register.
- **q**: cnt register.
- **y**: lcb register.
- **z**: scb register.
- **a**: cnt + lcb + scb register.
- **c**: cr0—cr15 register.
- **b**: cp1 registers.
- **f**: cp2 registers.
- **i**: cp3 registers.
- **j**: cp1 + cp2 + cp3 registers.
- **I**: High 16-bit constant (32-bit constant with 16 LSBs zero).
J Unsigned 5 bit integer (in the range 0 to 31).
K Unsigned 16 bit integer (in the range 0 to 65535).
L Signed 16 bit integer (in the range $-32768$ to $32767$).
M Unsigned 14 bit integer (in the range 0 to 16383).
N Signed 14 bit integer (in the range $-8192$ to $8191$).
Z Any SYMBOL_REF.

Xstormy16—‘config/stormy16/stormy16.h’

a Register r0.
b Register r1.
c Register r2.
d Register r8.
e Registers r0 through r7.
t Registers r0 and r1.
y The carry register.
z Registers r8 and r9.
I A constant between 0 and 3 inclusive.
J A constant that has exactly one bit set.
K A constant that has exactly one bit clear.
L A constant between 0 and 255 inclusive.
M A constant between $-255$ and 0 inclusive.
N A constant between $-3$ and 0 inclusive.
O A constant between 1 and 4 inclusive.
P A constant between $-4$ and $-1$ inclusive.
Q A memory reference that is a stack push.
R A memory reference that is a stack pop.
S A memory reference that refers to a constant address of known value.
T The register indicated by Rx (not implemented yet).
U A constant that is not between 2 and 15 inclusive.
Z The constant 0.

Xtensa—‘config/xtensa/constraints.md’

a General-purpose 32-bit register
b One-bit boolean register
A  MAC16 40-bit accumulator register
I  Signed 12-bit integer constant, for use in MOVI instructions
J  Signed 8-bit integer constant, for use in ADDI instructions
K  Integer constant valid for BccI instructions
L  Unsigned constant valid for BccUI instructions

16.8.6 Disable insn alternatives using the enabled attribute

The enabled insn attribute may be used to disable certain insn alternatives for machinespecific reasons. This is useful when adding new instructions to an existing pattern which are only available for certain cpu architecture levels as specified with the -march= option.

If an insn alternative is disabled, then it will never be used. The compiler treats the
constraints for the disabled alternative as unsatisfiable.

In order to make use of the enabled attribute a back end has to add in the machine
description files:

1. A definition of the enabled insn attribute. The attribute is defined as usual using the define_attr command. This definition should be based on other insn attributes and/or target flags. The enabled attribute is a numeric attribute and should evaluate to (const_int 1) for an enabled alternative and to (const_int 0) otherwise.

2. A definition of another insn attribute used to describe for what reason an insn alternative might be available or not. E.g. cpu_facility as in the example below.

3. An assignment for the second attribute to each insn definition combining instructions which are not all available under the same circumstances. (Note: It obviously only makes sense for definitions with more than one alternative. Otherwise the insn pattern should be disabled or enabled using the insn condition.)

E.g. the following two patterns could easily be merged using the enabled attribute:

```c
(define_insn "*movdi_old"
  [(set (match_operand:DI 0 "register_operand" "=d")
       (match_operand:DI 1 "register_operand" " d"))]
  "!TARGET_NEW"
  "lgr %0,%1")

(define_insn "*movdi_new"
  [(set (match_operand:DI 0 "register_operand" "=d,f,d")
       (match_operand:DI 1 "register_operand" " d,d,f"))]
  "TARGET_NEW"
  "@"
  lgr  %0,%1
  ldgr %0,%1
  lgdr %0,%1")
```
to:

```c
(define_insn "*movdi_combined"
  [(set (match_operand:DI 0 "register_operand" "=d,f,d")
       (match_operand:DI 1 "register_operand" " d,d,f"))]
  ""
```

"@lgr %0,%1
ldgr %0,%1
lgdr %0,%1"

[set_attr "cpu_facility" "*,new,new"]

with the enabled attribute defined like this:

(define_attr "cpu_facility" "standard,new" (const_string "standard"))

(define_attr "enabled" ""
  (cond [(eq_attr "cpu_facility" "standard") (const_int 1)
        (and (eq_attr "cpu_facility" "new")
         (ne (symbol_ref "TARGET_NEW") (const_int 0)))
      (const_int 1)]
      (const_int 0)))

### 16.8.7 Defining Machine-Specific Constraints

Machine-specific constraints fall into two categories: register and non-register constraints. Within the latter category, constraints which allow subsets of all possible memory or address operands should be specially marked, to give reload more information.

Machine-specific constraints can be given names of arbitrary length, but they must be entirely composed of letters, digits, underscores ('_'), and angle brackets ('< >'). Like C identifiers, they must begin with a letter or underscore.

In order to avoid ambiguity in operand constraint strings, no constraint can have a name that begins with any other constraint’s name. For example, if \( x \) is defined as a constraint name, \( xy \) may not be, and vice versa. As a consequence of this rule, no constraint may begin with one of the generic constraint letters: ‘E F V X g i m n o p r s’.

Register constraints correspond directly to register classes. See Section 17.8 [Register Classes], page 425. There is thus not much flexibility in their definitions.

**define_register_constraint name regclass docstring**  
All three arguments are string constants. *name* is the name of the constraint, as it will appear in \( \text{match_operand} \) expressions. If *name* is a multi-letter constraint its length shall be the same for all constraints starting with the same letter. *regclass* can be either the name of the corresponding register class (see Section 17.8 [Register Classes], page 425), or a C expression which evaluates to the appropriate register class. If it is an expression, it must have no side effects, and it cannot look at the operand. The usual use of expressions is to map some register constraints to \text{NO_REGS} when the register class is not available on a given subarchitecture.  
*docstring* is a sentence documenting the meaning of the constraint. Docstrings are explained further below.

Non-register constraints are more like predicates: the constraint definition gives a Boolean expression which indicates whether the constraint matches.

**define_constraint name docstring exp**  
The *name* and *docstring* arguments are the same as for **define_register_constraint**, but note that the docstring comes immediately after the name for
these expressions. exp is an RTL expression, obeying the same rules as the RTL expressions in predicate definitions. See Section 16.7.2 [Defining Predicates], page 300, for details. If it evaluates true, the constraint matches; if it evaluates false, it doesn’t. Constraint expressions should indicate which RTL codes they might match, just like predicate expressions.

**match_test** C expressions have access to the following variables:

- **op** The RTL object defining the operand.
- **mode** The machine mode of op.
- **ival** ‘INTVAL (op)’, if op is a const_int.
- **hval** ‘CONST_DOUBLE_HIGH (op)’, if op is an integer const_double.
- **lval** ‘CONST_DOUBLE_LOW (op)’, if op is an integer const_double.
- **rval** ‘CONST_DOUBLE_REAL_VALUE (op)’, if op is a floating-point const_double.

The *val variables should only be used once another piece of the expression has verified that op is the appropriate kind of RTL object.

Most non-register constraints should be defined with **define_constraint**. The remaining two definition expressions are only appropriate for constraints that should be handled specially by **reload** if they fail to match.

**define_memory_constraint** name docstring exp  
[MD Expression]
Use this expression for constraints that match a subset of all memory operands: that is, **reload** can make them match by converting the operand to the form ‘(mem (reg X))’, where X is a base register (from the register class specified by BASE_REG_CLASS, see Section 17.8 [Register Classes], page 425).

For example, on the S/390, some instructions do not accept arbitrary memory references, but only those that do not make use of an index register. The constraint letter ‘Q’ is defined to represent a memory address of this type. If ‘Q’ is defined with **define_memory_constraint**, a ‘Q’ constraint can handle any memory operand, because **reload** knows it can simply copy the memory address into a base register if required. This is analogous to the way an ‘o’ constraint can handle any memory operand.

The syntax and semantics are otherwise identical to **define_constraint**.

**define_address_constraint** name docstring exp  
[MD Expression]
Use this expression for constraints that match a subset of all address operands: that is, **reload** can make the constraint match by converting the operand to the form ‘(reg X)’, again with X a base register.

Constraints defined with **define_address_constraint** can only be used with the **address_operand** predicate, or machine-specific predicates that work the same way. They are treated analogously to the generic ‘p’ constraint.

The syntax and semantics are otherwise identical to **define_constraint**.
For historical reasons, names beginning with the letters ‘G H’ are reserved for constraints that match only `const doubles`, and names beginning with the letters ‘I J K L M N O P’ are reserved for constraints that match only `const ints`. This may change in the future. For the time being, constraints with these names must be written in a stylized form, so that `genpreds` can tell you did it correctly:

```lisp
(define_constraint "[GHIJKLMNOP]...
"doc...")
(and (match_code "const_int") ; const_double for G/H
    condition...)) ; usually a match_test
```

It is fine to use names beginning with other letters for constraints that match `const doubles` or `const ints`.

Each docstring in a constraint definition should be one or more complete sentences, marked up in Texinfo format. They are currently unused. In the future they will be copied into the GCC manual, in Section 16.8.5 [Machine Constraints], page 308, replacing the hand-maintained tables currently found in that section. Also, in the future the compiler may use this to give more helpful diagnostics when poor choice of `asm` constraints causes a reload failure.

If you put the pseudo-Texinfo directive `@internal` at the beginning of a docstring, then (in the future) it will appear only in the internals manual’s version of the machine-specific constraint tables. Use this for constraints that should not appear in `asm` statements.

### 16.8.8 Testing constraints from C

It is occasionally useful to test a constraint from C code rather than implicitly via the constraint string in a `match_operand`. The generated file ‘tm_p.h’ declares a few interfaces for working with machine-specific constraints. None of these interfaces work with the generic constraints described in Section 16.8.1 [Simple Constraints], page 302. This may change in the future.

**Warning:** ‘tm_p.h’ may declare other functions that operate on constraints, besides the ones documented here. Do not use those functions from machine-dependent code. They exist to implement the old constraint interface that machine-independent components of the compiler still expect. They will change or disappear in the future.

Some valid constraint names are not valid C identifiers, so there is a mangling scheme for referring to them from C. Constraint names that do not contain angle brackets or underscores are left unchanged. Underscores are doubled, each ‘<’ is replaced with ‘_l’, and each ‘>’ with ‘_g’. Here are some examples:

<table>
<thead>
<tr>
<th>Original</th>
<th>Mangled</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>P42x</td>
<td>P42x</td>
</tr>
<tr>
<td>P4_x</td>
<td>P4__x</td>
</tr>
<tr>
<td>P4&gt;</td>
<td>P4_gx</td>
</tr>
<tr>
<td>P4&gt;&gt;</td>
<td>P4_g_g</td>
</tr>
<tr>
<td>P4&gt;_</td>
<td>P4__g_g</td>
</tr>
</tbody>
</table>

Throughout this section, the variable `c` is either a constraint in the abstract sense, or a constant from `enum constraint_num`; the variable `m` is a mangled constraint name (usually as part of a larger identifier).
constraint_num

For each machine-specific constraint, there is a corresponding enumeration constant: ‘CONSTRAINT_’ plus the mangled name of the constraint. Functions that take an enum constraint_num as an argument expect one of these constants.

Machine-independent constraints do not have associated constants. This may change in the future.

inline bool satisfies_constraint_m (rtx exp)     [Function]
For each machine-specific, non-register constraint m, there is one of these functions; it returns true if exp satisfies the constraint. These functions are only visible if ‘rtl.h’ was included before ‘tm_p.h’.

bool constraint_satisfied_p (rtx exp, enum constraint_num c)     [Function]
Like the satisfies_constraint_m functions, but the constraint to test is given as an argument, c. If c specifies a register constraint, this function will always return false.

enum reg_class regclass_for_constraint (enum constraint_num c)     [Function]
Returns the register class associated with c. If c is not a register constraint, or those registers are not available for the currently selected subtarget, returns NO_REGS.

Here is an example use of satisfies_constraint_m. In peephole optimizations (see Section 16.18 [Peephole Definitions], page 368), operand constraint strings are ignored, so if there are relevant constraints, they must be tested in the C condition. In the example, the optimization is applied if operand 2 does not satisfy the ‘K’ constraint. (This is a simplified version of a peephole definition from the i386 machine description.)

(define_peephole2
 [(match_scratch:SI 3 "r")
  (set (match_operand:SI 0 "register_operand" ""
       (mult:SI (match_operand:SI 1 "memory_operand" ""
               (match_operand:SI 2 "immediate_operand" ""))))))]

"!satisfies_constraint_K (operands[2])"

[(set (match_dup 3) (match_dup 1))
  (set (match_dup 0) (mult:SI (match_dup 3) (match_dup 2)))]

"

16.9 Standard Pattern Names For Generation

Here is a table of the instruction names that are meaningful in the RTL generation pass of the compiler. Giving one of these names to an instruction pattern tells the RTL generation pass that it can use the pattern to accomplish a certain task.

‘movm’     Here m stands for a two-letter machine mode name, in lowercase. This instruction pattern moves data with that machine mode from operand 1 to operand 0. For example, ‘movsi’ moves full-word data.

    If operand 0 is a subreg with mode m of a register whose own mode is wider than m, the effect of this instruction is to store the specified value in the part of the register that corresponds to mode m. Bits outside of m, but which
are within the same target word as the subreg are undefined. Bits which are outside the target word are left unchanged.

This class of patterns is special in several ways. First of all, each of these names up to and including full word size must be defined, because there is no other way to copy a datum from one place to another. If there are patterns accepting operands in larger modes, movm must be defined for integer modes of those sizes.

Second, these patterns are not used solely in the RTL generation pass. Even the reload pass can generate move insns to copy values from stack slots into temporary registers. When it does so, one of the operands is a hard register and the other is an operand that can need to be reloaded into a register.

Therefore, when given such a pair of operands, the pattern must generate RTL which needs no reloading and needs no temporary registers—no registers other than the operands. For example, if you support the pattern with a define_expand, then in such a case the define_expand mustn’t call force_reg or any other such function which might generate new pseudo registers.

This requirement exists even for subword modes on a RISC machine where fetching those modes from memory normally requires several insns and some temporary registers.

During reload a memory reference with an invalid address may be passed as an operand. Such an address will be replaced with a valid address later in the reload pass. In this case, nothing may be done with the address except to use it as it stands. If it is copied, it will not be replaced with a valid address. No attempt should be made to make such an address into a valid address and no routine (such as change_address) that will do so may be called. Note that general_operand will fail when applied to such an address.

The global variable reload_in_progress (which must be explicitly declared if required) can be used to determine whether such special handling is required.

The variety of operands that have reloads depends on the rest of the machine description, but typically on a RISC machine these can only be pseudo registers that did not get hard registers, while on other machines explicit memory references will get optional reloads.

If a scratch register is required to move an object to or from memory, it can be allocated using gen_reg_rtx prior to life analysis.

If there are cases which need scratch registers during or after reload, you must provide an appropriate secondary_reload target hook.

The macro can_create_pseudo_p can be used to determine if it is unsafe to create new pseudo registers. If this variable is nonzero, then it is unsafe to call gen_reg_rtx to allocate a new pseudo.

The constraints on a movm must permit moving any hard register to any other hard register provided that HARD_REGNO_MODE_OK permits mode \( m \) in both registers and TARGET_REGISTER_MOVE_COST applied to their classes returns a value of 2.

It is obligatory to support floating point movm instructions into and out of any registers that can hold fixed point values, because unions and structures (which
have modes \texttt{SImode} or \texttt{DImode}) can be in those registers and they may have floating point members.

There may also be a need to support fixed point ‘\texttt{movm}’ instructions in and out of floating point registers. Unfortunately, I have forgotten why this was so, and I don’t know whether it is still true. If \texttt{HARD_REGNO_MODE_OK} rejects fixed point values in floating point registers, then the constraints of the fixed point ‘\texttt{movm}’ instructions must be designed to avoid ever trying to reload into a floating point register.

‘\texttt{reload_inm}\texttt{}\texttt{m}’

‘\texttt{reload_outm}\texttt{}\texttt{m}’

These named patterns have been obsoleted by the target hook \texttt{secondary_reload}.

Like ‘\texttt{movm}’, but used when a scratch register is required to move between operand 0 and operand 1. Operand 2 describes the scratch register. See the discussion of the \texttt{SECONDARY_RELOAD_CLASS} macro in see Section 17.8 [Register Classes], page 425.

There are special restrictions on the form of the \texttt{match_operands} used in these patterns. First, only the predicate for the reload operand is examined, i.e., \texttt{reload_in} examines operand 1, but not the predicates for operand 0 or 2. Second, there may be only one alternative in the constraints. Third, only a single register class letter may be used for the constraint; subsequent constraint letters are ignored. As a special exception, an empty constraint string matches the \texttt{ALL_REGS} register class. This may relieve ports of the burden of defining an \texttt{ALL_REGS} constraint letter just for these patterns.

‘\texttt{movstrictm}\texttt{}\texttt{m}’

Like ‘\texttt{movm}’ except that if operand 0 is a \texttt{subreg} with mode \texttt{m} of a register whose natural mode is wider, the ‘\texttt{movstrictm}\texttt{m}’ instruction is guaranteed not to alter any of the register except the part which belongs to mode \texttt{m}.

‘\texttt{movmisalignm}\texttt{}\texttt{m}’

This variant of a move pattern is designed to load or store a value from a memory address that is not naturally aligned for its mode. For a store, the memory will be in operand 0; for a load, the memory will be in operand 1. The other operand is guaranteed not to be a memory, so that it's easy to tell whether this is a load or store.

This pattern is used by the autovectorizer, and when expanding a \texttt{MISALIGNED INDIRECT_REF} expression.

‘\texttt{load_multiple}\texttt{}\texttt{m}’

Load several consecutive memory locations into consecutive registers. Operand 0 is the first of the consecutive registers, operand 1 is the first memory location, and operand 2 is a constant: the number of consecutive registers.

Define this only if the target machine really has such an instruction; do not define this if the most efficient way of loading consecutive registers from memory is to do them one at a time.
On some machines, there are restrictions as to which consecutive registers can be stored into memory, such as particular starting or ending register numbers or only a range of valid counts. For those machines, use a define_expand (see Section 16.15 [Expander Definitions], page 361) and make the pattern fail if the restrictions are not met.

Write the generated insn as a parallel with elements being a set of one register from the appropriate memory location (you may also need use or clobber elements). Use a match_parallel (see Section 16.4 [RTL Template], page 291) to recognize the insn. See ‘rs6000.md’ for examples of the use of this insn pattern.

'store_multiple'
Similar to 'load_multiple', but store several consecutive registers into consecutive memory locations. Operand 0 is the first of the consecutive memory locations, operand 1 is the first register, and operand 2 is a constant: the number of consecutive registers.

'vec_setm'
Set given field in the vector value. Operand 0 is the vector to modify, operand 1 is new value of field and operand 2 specify the field index.

'vec_extractm'
Extract given field from the vector value. Operand 1 is the vector, operand 2 specify field index and operand 0 place to store value into.

'vec_extract_evenm'
Extract even elements from the input vectors (operand 1 and operand 2). The even elements of operand 2 are concatenated to the even elements of operand 1 in their original order. The result is stored in operand 0. The output and input vectors should have the same modes.

'vec_extract_oddm'
Extract odd elements from the input vectors (operand 1 and operand 2). The odd elements of operand 2 are concatenated to the odd elements of operand 1 in their original order. The result is stored in operand 0. The output and input vectors should have the same modes.

'vec_interleave_highm'
Merge high elements of the two input vectors into the output vector. The output and input vectors should have the same modes (N elements). The high N/2 elements of the first input vector are interleaved with the high N/2 elements of the second input vector.

'vec_interleave_lowm'
Merge low elements of the two input vectors into the output vector. The output and input vectors should have the same modes (N elements). The low N/2 elements of the first input vector are interleaved with the low N/2 elements of the second input vector.

'vec_initm'
Initialize the vector to given values. Operand 0 is the vector to initialize and operand 1 is parallel containing values for individual fields.
Output a push instruction. Operand 0 is value to push. Used only when PUSH_ROUNDING is defined. For historical reason, this pattern may be missing and in such case an mov expander is used instead, with a MEM expression forming the push operation. The mov expander method is deprecated.

Add operand 2 and operand 1, storing the result in operand 0. All operands must have mode m. This can be used even on two-address machines, by means of constraints requiring operands 1 and 0 to be the same location.

Similar, for other arithmetic operations.

Multiply operand 2 and operand 1, then add operand 3, storing the result in operand 0. All operands must have mode m. This pattern is used to implement the fma, fmaf, and fmal builtin functions from the ISO C99 standard. The fma operation may produce different results than doing the multiply followed by the add if the machine does not perform a rounding step between the operations.

Like fmam4, except operand 3 subtracted from the product instead of added to the product. This is represented in the rtl as

(fma:m (neg:m op1) op2 op3)

Like fmam4 except that the intermediate product is negated before being added to operand 3. This is represented in the rtl as

(fma:m op1 op2 (neg:m op3))

Like fmsm4 except that the intermediate product is negated before subtracting operand 3. This is represented in the rtl as

(fma:m (neg:m op1) op2 (neg:m op3))

Signed minimum and maximum operations. When used with floating point, if both operands are zeros, or if either operand is NaN, then it is unspecified which of the two operands is returned as the result.

Find the signed minimum/maximum of the elements of a vector. The vector is operand 1, and the scalar result is stored in the least significant bits of operand 0 (also a vector). The output and input vector should have the same modes.

Find the unsigned minimum/maximum of the elements of a vector. The vector is operand 1, and the scalar result is stored in the least significant bits of operand 0 (also a vector). The output and input vector should have the same modes.
‘reduc_splus_m’
Compute the sum of the signed elements of a vector. The vector is operand 1, and the scalar result is stored in the least significant bits of operand 0 (also a vector). The output and input vector should have the same modes.

‘reduc_uplus_m’
Compute the sum of the unsigned elements of a vector. The vector is operand 1, and the scalar result is stored in the least significant bits of operand 0 (also a vector). The output and input vector should have the same modes.

‘sdot_prodm’
‘udot_prodm’
Compute the sum of the products of two signed/unsigned elements. Operand 1 and operand 2 are of the same mode. Their product, which is of a wider mode, is computed and added to operand 3. Operand 3 is of a mode equal or wider than the mode of the product. The result is placed in operand 0, which is of the same mode as operand 3.

‘ssum_widenm3’
‘usum_widenm3’
Operands 0 and 2 are of the same mode, which is wider than the mode of operand 1. Add operand 1 to operand 2 and place the widened result in operand 0. (This is used express accumulation of elements into an accumulator of a wider mode.)

‘vec_shl_m’, ‘vec_shr_m’
Whole vector left/right shift in bits. Operand 1 is a vector to be shifted. Operand 2 is an integer shift amount in bits. Operand 0 is where the resulting shifted vector is stored. The output and input vectors should have the same modes.

‘vec_pack_trunc_m’
Narrow (demote) and merge the elements of two vectors. Operands 1 and 2 are vectors of the same mode having N integral or floating point elements of size S. Operand 0 is the resulting vector in which 2*N elements of size N/2 are concatenated after narrowing them down using truncation.

‘vec_pack_ssat_m’, ‘vec_pack_usat_m’
Narrow (demote) and merge the elements of two vectors. Operands 1 and 2 are vectors of the same mode having N integral elements of size S. Operand 0 is the resulting vector in which the elements of the two input vectors are concatenated after narrowing them down using signed/unsigned saturating arithmetic.

‘vec_pack_sfix_trunc_m’, ‘vec_pack_ufix_trunc_m’
Narrow, convert to signed/unsigned integral type and merge the elements of two vectors. Operands 1 and 2 are vectors of the same mode having N floating point elements of size S. Operand 0 is the resulting vector in which 2*N elements of size N/2 are concatenated.

‘vec_unpacks_hi_m’, ‘vec_unpacks_lo_m’
Extract and widen (promote) the high/low part of a vector of signed integral or floating point elements. The input vector (operand 1) has N elements of size S.
Widen (promote) the high/low elements of the vector using signed or floating point extension and place the resulting N/2 values of size 2*S in the output vector (operand 0).

‘vec_unpacku_hi_m’, ‘vec_unpacku_lo_m’
Extract and widen (promote) the high/low part of a vector of unsigned integral elements. The input vector (operand 1) has N elements of size S. Widen (promote) the high/low elements of the vector using zero extension and place the resulting N/2 values of size 2*S in the output vector (operand 0).

‘vec_unpacks_float_hi_m’, ‘vec_unpacks_float_lo_m’
‘vec_unpacku_float_hi_m’, ‘vec_unpacku_float_lo_m’
Extract, convert to floating point type and widen the high/low part of a vector of signed/unsigned integral elements. The input vector (operand 1) has N elements of size S. Convert the high/low elements of the vector using floating point conversion and place the resulting N/2 values of size 2*S in the output vector (operand 0).

‘vec_widen_umult_hi_m’, ‘vec_widen_umult_lo_m’
‘vec_widen_smult_hi_m’, ‘vec_widen_smult_lo_m’
Signed/Unsigned widening multiplication. The two inputs (operands 1 and 2) are vectors with N signed/unsigned elements of size S. Multiply the high/low elements of the two vectors, and put the N/2 products of size 2*S in the output vector (operand 0).

‘mulhis13’
Multiply operands 1 and 2, which have mode HImode, and store a SImode product in operand 0.

‘mulqih13’, ‘mulsidi3’
Similar widening-multiplication instructions of other widths.

‘umulqih13’, ‘umulhis13’, ‘umulsidi3’
Similar widening-multiplication instructions that do unsigned multiplication.

‘usmulqih13’, ‘usmulhis13’, ‘usmulsidi3’
Similar widening-multiplication instructions that interpret the first operand as unsigned and the second operand as signed, then do a signed multiplication.

‘smulm3_highpart’
Perform a signed multiplication of operands 1 and 2, which have mode m, and store the most significant half of the product in operand 0. The least significant half of the product is discarded.

‘umulm3_highpart’
Similar, but the multiplication is unsigned.

‘maddmn4’
Multiply operands 1 and 2, sign-extend them to mode n, add operand 3, and store the result in operand 0. Operands 1 and 2 have mode m and operands 0 and 3 have mode n. Both modes must be integer or fixed-point modes and n must be twice the size of m.

In other words, maddmn4 is like mulmn3 except that it also adds operand 3. These instructions are not allowed to FAIL.
Like \texttt{madd}_{mn}^4, but zero-extend the multiplication operands instead of sign-extend them.

Like \texttt{madd}_{mn}^4, but all involved operations must be signed-saturating.

Like \texttt{umadd}_{mn}^4, but all involved operations must be unsigned-saturating.

Multiply operands 1 and 2, sign-extend them to mode \( n \), subtract the result from operand 3, and store the result in operand 0. Operands 1 and 2 have mode \( m \) and operands 0 and 3 have mode \( n \). Both modes must be integer or fixed-point modes and \( n \) must be twice the size of \( m \).

In other words, \texttt{msub}_{mn}^4 is like \texttt{mul}_{mn}^3 except that it also subtracts the result from operand 3.

These instructions are not allowed to \texttt{FAIL}.

Like \texttt{msub}_{mn}^4, but zero-extend the multiplication operands instead of sign-extend them.

Like \texttt{msub}_{mn}^4, but all involved operations must be signed-saturating.

Like \texttt{umsub}_{mn}^4, but all involved operations must be unsigned-saturating.

Signed division that produces both a quotient and a remainder. Operand 1 is divided by operand 2 to produce a quotient stored in operand 0 and a remainder stored in operand 3.

For machines with an instruction that produces both a quotient and a remainder, provide a pattern for \texttt{divmod}_{mn}^4 but do not provide patterns for \texttt{div}_{mn}^3 and \texttt{mod}_{mn}^3. This allows optimization in the relatively common case when both the quotient and remainder are computed.

If an instruction that just produces a quotient or just a remainder exists and is more efficient than the instruction that produces both, write the output routine of \texttt{divmod}_{mn}^4 to call \texttt{find\_reg\_note} and look for a \texttt{REG\_unused} note on the quotient or remainder and generate the appropriate instruction.

Similar, but does unsigned division.

Arithmetic-shift operand 1 left by a number of bits specified by operand 2, and store the result in operand 0. Here \( m \) is the mode of operand 0 and operand 1; operand 2's mode is specified by the instruction pattern, and the compiler will convert the operand to that mode before generating the instruction. The meaning of out-of-range shift counts can optionally be specified by \texttt{TARGET\_SHIFT\_TRUNCATION\_MASK}. See \cite{TARGET\_SHIFT\_TRUNCATION\_MASK}, page 544. Operand 2 is always a scalar type.
'ashr3', 'lshr3', 'rotl3', 'rotr3'
   Other shift and rotate instructions, analogous to the ashlm3 instructions.
   Operand 2 is always a scalar type.

'vashl3', 'vashr3', 'vlshr3', 'vrotl3', 'vrotr3'
   Vector shift and rotate instructions that take vectors as operand 2 instead of a
   scalar type.

'neg2', 'ssneg2', 'usneg2'
   Negate operand 1 and store the result in operand 0.

'absm2'
   Store the absolute value of operand 1 into operand 0.

'sqrtm2'
   Store the square root of operand 1 into operand 0.
   The sqrt built-in function of C always uses the mode which corresponds to
   the C data type double and the sqrtf built-in function uses the mode which
   corresponds to the C data type float.

'fmodm3'
   Store the remainder of dividing operand 1 by operand 2 into operand 0, rounded
   towards zero to an integer.
   The fmod built-in function of C always uses the mode which corresponds to
   the C data type double and the fmodf built-in function uses the mode which
   corresponds to the C data type float.

'remaindern3'
   Store the remainder of dividing operand 1 by operand 2 into operand 0, rounded
   to the nearest integer.
   The remainder built-in function of C always uses the mode which corresponds
   to the C data type double and the remaindernf built-in function uses the mode
   which corresponds to the C data type float.

'cosm2'
   Store the cosine of operand 1 into operand 0.
   The cos built-in function of C always uses the mode which corresponds to
   the C data type double and the cosf built-in function uses the mode which
   corresponds to the C data type float.

'sinm2'
   Store the sine of operand 1 into operand 0.
   The sin built-in function of C always uses the mode which corresponds to
   the C data type double and the sinf built-in function uses the mode which
   corresponds to the C data type float.

'expm2'
   Store the exponential of operand 1 into operand 0.
   The exp built-in function of C always uses the mode which corresponds to
   the C data type double and the expf built-in function uses the mode which
   corresponds to the C data type float.

'logm2'
   Store the natural logarithm of operand 1 into operand 0.
   The log built-in function of C always uses the mode which corresponds to
   the C data type double and the logf built-in function uses the mode which
   corresponds to the C data type float.
'powm3' Store the value of operand 1 raised to the exponent operand 2 into operand 0. The pow built-in function of C always uses the mode which corresponds to the C data type double and the powf built-in function uses the mode which corresponds to the C data type float.

'atan2m3' Store the arc tangent (inverse tangent) of operand 1 divided by operand 2 into operand 0, using the signs of both arguments to determine the quadrant of the result.

The atan2 built-in function of C always uses the mode which corresponds to the C data type double and the atan2f built-in function uses the mode which corresponds to the C data type float.

'floorm2' Store the largest integral value not greater than argument.

The floor built-in function of C always uses the mode which corresponds to the C data type double and the floorf built-in function uses the mode which corresponds to the C data type float.

'btruncm2' Store the argument rounded to integer towards zero.

The trunc built-in function of C always uses the mode which corresponds to the C data type double and the truncf built-in function uses the mode which corresponds to the C data type float.

'roundm2' Store the argument rounded to integer away from zero.

The round built-in function of C always uses the mode which corresponds to the C data type double and the roundf built-in function uses the mode which corresponds to the C data type float.

'ceilm2' Store the argument rounded to integer away from zero.

The ceil built-in function of C always uses the mode which corresponds to the C data type double and the ceilf built-in function uses the mode which corresponds to the C data type float.

'nearbyintm2' Store the argument rounded according to the default rounding mode

The nearbyint built-in function of C always uses the mode which corresponds to the C data type double and the nearbyintf built-in function uses the mode which corresponds to the C data type float.

'rintm2' Store the argument rounded according to the default rounding mode and raise the inexact exception when the result differs in value from the argument

The rint built-in function of C always uses the mode which corresponds to the C data type double and the rintf built-in function uses the mode which corresponds to the C data type float.

'lrintmn2' Convert operand 1 (valid for floating point mode m) to fixed point mode n as a signed number according to the current rounding mode and store in operand 0 (which has mode n).
'lroundmn2'
Convert operand 1 (valid for floating point mode \( m \)) to fixed point mode \( n \) as a signed number rounding to nearest and away from zero and store in operand 0 (which has mode \( n \)).

'lfloormn2'
Convert operand 1 (valid for floating point mode \( m \)) to fixed point mode \( n \) as a signed number rounding down and store in operand 0 (which has mode \( n \)).

'lceilmn2'
Convert operand 1 (valid for floating point mode \( m \)) to fixed point mode \( n \) as a signed number rounding up and store in operand 0 (which has mode \( n \)).

'copysignm3'
Store a value with the magnitude of operand 1 and the sign of operand 2 into operand 0. The \texttt{copysign} built-in function of C always uses the mode which corresponds to the C data type \texttt{double} and the \texttt{copysignf} built-in function uses the mode which corresponds to the C data type \texttt{float}.

'ffsm2'
Store into operand 0 one plus the index of the least significant 1-bit of operand 1. If operand 1 is zero, store zero. \( m \) is the mode of operand 0; operand 1’s mode is specified by the instruction pattern, and the compiler will convert the operand to that mode before generating the instruction. The \texttt{ffs} built-in function of C always uses the mode which corresponds to the C data type \texttt{int}.

'clzm2'
Store into operand 0 the number of leading 0-bits in \( x \), starting at the most significant bit position. If \( x \) is 0, the \texttt{CLZ_DEFINED_VALUE_AT_ZERO} (see Section 17.31 [Misc], page 542) macro defines if the result is undefined or has a useful value. \( m \) is the mode of operand 0; operand 1’s mode is specified by the instruction pattern, and the compiler will convert the operand to that mode before generating the instruction.

'ctzm2'
Store into operand 0 the number of trailing 0-bits in \( x \), starting at the least significant bit position. If \( x \) is 0, the \texttt{CTZ_DEFINED_VALUE_AT_ZERO} (see Section 17.31 [Misc], page 542) macro defines if the result is undefined or has a useful value. \( m \) is the mode of operand 0; operand 1’s mode is specified by the instruction pattern, and the compiler will convert the operand to that mode before generating the instruction.

'popcountm2'
Store into operand 0 the number of 1-bits in \( x \). \( m \) is the mode of operand 0; operand 1’s mode is specified by the instruction pattern, and the compiler will convert the operand to that mode before generating the instruction.

'paritym2'
Store into operand 0 the parity of \( x \), i.e. the number of 1-bits in \( x \) modulo 2. \( m \) is the mode of operand 0; operand 1’s mode is specified by the instruction pattern, and the compiler will convert the operand to that mode before generating the instruction.
'one_cmplm2'
Store the bitwise-complement of operand 1 into operand 0.

'movmem'
Block move instruction. The destination and source blocks of memory are the first two operands, and both are mem:BLKs with an address in mode Pmode. The number of bytes to move is the third operand, in mode m. Usually, you specify word_mode for m. However, if you can generate better code knowing the range of valid lengths is smaller than those representable in a full word, you should provide a pattern with a mode corresponding to the range of values you can handle efficiently (e.g., QImode for values in the range 0–127; note we avoid numbers that appear negative) and also a pattern with word_mode.

The fourth operand is the known shared alignment of the source and destination, in the form of a const_int rtx. Thus, if the compiler knows that both source and destination are word-aligned, it may provide the value 4 for this operand.

Optional operands 5 and 6 specify expected alignment and size of block respectively. The expected alignment differs from alignment in operand 4 in a way that the blocks are not required to be aligned according to it in all cases. This expected alignment is also in bytes, just like operand 4. Expected size, when unknown, is set to (const_int -1).

Descriptions of multiple movmem patterns can only be beneficial if the patterns for smaller modes have fewer restrictions on their first, second and fourth operands. Note that the mode m in movmem does not impose any restriction on the mode of individually moved data units in the block.

These patterns need not give special consideration to the possibility that the source and destination strings might overlap.

'movstr'
String copy instruction, with stpcpy semantics. Operand 0 is an output operand in mode Pmode. The addresses of the destination and source strings are operands 1 and 2, and both are mem:BLKs with addresses in mode Pmode. The execution of the expansion of this pattern should store in operand 0 the address in which the NUL terminator was stored in the destination string.

'setmemm'
Block set instruction. The destination string is the first operand, given as a mem:BLK whose address is in mode Pmode. The number of bytes to set is the second operand, in mode m. The value to initialize the memory with is the third operand. Targets that only support the clearing of memory should reject any value that is not the constant 0. See 'movmem' for a discussion of the choice of mode.

The fourth operand is the known alignment of the destination, in the form of a const_int rtx. Thus, if the compiler knows that the destination is word-aligned, it may provide the value 4 for this operand.

Optional operands 5 and 6 specify expected alignment and size of block respectively. The expected alignment differs from alignment in operand 4 in a way that the blocks are not required to be aligned according to it in all cases. This expected alignment is also in bytes, just like operand 4. Expected size, when unknown, is set to (const_int -1).

The use for multiple setmemm is as for movmem.
'cmpstrnm'
String compare instruction, with five operands. Operand 0 is the output; it has mode m. The remaining four operands are like the operands of 'movmemmm'. The two memory blocks specified are compared byte by byte in lexicographic order starting at the beginning of each string. The instruction is not allowed to prefetch more than one byte at a time since either string may end in the first byte and reading past that may access an invalid page or segment and cause a fault. The comparison terminates early if the fetched bytes are different or if they are equal to zero. The effect of the instruction is to store a value in operand 0 whose sign indicates the result of the comparison.

'cmpstrm'
String compare instruction, without known maximum length. Operand 0 is the output; it has mode m. The second and third operand are the blocks of memory to be compared; both are mem:BLK with an address in mode Pmode. The fourth operand is the known shared alignment of the source and destination, in the form of a const_int rtx. Thus, if the compiler knows that both source and destination are word-aligned, it may provide the value 4 for this operand. The two memory blocks specified are compared byte by byte in lexicographic order starting at the beginning of each string. The instruction is not allowed to prefetch more than one byte at a time since either string may end in the first byte and reading past that may access an invalid page or segment and cause a fault. The comparison will terminate when the fetched bytes are different or if they are equal to zero. The effect of the instruction is to store a value in operand 0 whose sign indicates the result of the comparison.

'cmpmemm'
Block compare instruction, with five operands like the operands of 'cmpstrm'. The two memory blocks specified are compared byte by byte in lexicographic order starting at the beginning of each block. Unlike 'cmpstrm' the instruction can prefetch any bytes in the two memory blocks. Also unlike 'cmpstrm' the comparison will not stop if both bytes are zero. The effect of the instruction is to store a value in operand 0 whose sign indicates the result of the comparison.

'strlenm'
Compute the length of a string, with three operands. Operand 0 is the result (of mode m), operand 1 is a mem referring to the first character of the string, operand 2 is the character to search for (normally zero), and operand 3 is a constant describing the known alignment of the beginning of the string.

'floatmn2'
Convert signed integer operand 1 (valid for fixed point mode m) to floating point mode n and store in operand 0 (which has mode n).

'floatunsmn2'
Convert unsigned integer operand 1 (valid for fixed point mode m) to floating point mode n and store in operand 0 (which has mode n).

'fixmn2'
Convert operand 1 (valid for floating point mode m) to fixed point mode n as a signed number and store in operand 0 (which has mode n). This instruction's result is defined only when the value of operand 1 is an integer. If the machine description defines this pattern, it also needs to define the ftrunc pattern.
‘fixuns\_mn2’
  Convert operand 1 (valid for floating point mode m) to fixed point mode n as an unsigned number and store in operand 0 (which has mode n). This instruction’s result is defined only when the value of operand 1 is an integer.

‘ftrunc\_m2’
  Convert operand 1 (valid for floating point mode m) to an integer value, still represented in floating point mode m, and store it in operand 0 (valid for floating point mode m).

‘fix\_trunc\_mn2’
  Like ‘fixmn2’ but works for any floating point value of mode m by converting the value to an integer.

‘fixuns\_trunc\_mn2’
  Like ‘fixunsmn2’ but works for any floating point value of mode m by converting the value to an integer.

‘trunc\_mn2’
  Truncate operand 1 (valid for mode m) to mode n and store in operand 0 (which has mode n). Both modes must be fixed point or both floating point.

‘extend\_mn2’
  Sign-extend operand 1 (valid for mode m) to mode n and store in operand 0 (which has mode n). Both modes must be fixed point or both floating point.

‘zero\_extend\_mn2’
  Zero-extend operand 1 (valid for mode m) to mode n and store in operand 0 (which has mode n). Both modes must be fixed point.

‘fract\_mn2’
  Convert operand 1 of mode m to mode n and store in operand 0 (which has mode n). Mode m and mode n could be fixed-point to fixed-point, signed integer to fixed-point, fixed-point to signed integer, floating-point to fixed-point, or fixed-point to floating-point. When overflows or underflows happen, the results are undefined.

‘satfract\_mn2’
  Convert operand 1 of mode m to mode n and store in operand 0 (which has mode n). Mode m and mode n could be fixed-point to fixed-point, signed integer to fixed-point, or floating-point to fixed-point. When overflows or underflows happen, the instruction saturates the results to the maximum or the minimum.

‘fractuns\_mn2’
  Convert unsigned integer operand 1 of mode m to fixed-point mode n and store in operand 0 (which has mode n). When overflows or underflows happen, the results are undefined.

‘satfractuns\_mn2’
  Convert unsigned integer operand 1 of mode m to fixed-point mode n and store in operand 0 (which has mode n). When overflows or underflows happen, the instruction saturates the results to the maximum or the minimum.
The RTL generation pass generates this instruction only with constants for operands 2 and 3 and the constant is never zero for operand 2.

The bit-field value is sign-extended to a full word integer before it is stored in operand 0.

Like `extv` except that the bit-field value is zero-extended.

Store operand 3 (which must be valid for `word_mode`) into a bit-field in operand 0, where operand 1 specifies the width in bits and operand 2 the starting bit. Operand 0 may have mode `byte_mode` or `word_mode`; often `word_mode` is allowed only for registers. Operands 1 and 2 must be valid for `word_mode`.

The RTL generation pass generates this instruction only with constants for operands 1 and 2 and the constant is never zero for operand 1.

Conditionally move operand 2 or operand 3 into operand 0 according to the comparison in operand 1. If the comparison is true, operand 2 is moved into operand 0, otherwise operand 3 is moved.

The mode of the operands being compared need not be the same as the operands being moved. Some machines, sparc64 for example, have instructions that conditionally move an integer value based on the floating point condition codes and vice versa.

If the machine does not have conditional move instructions, do not define these patterns.

Similar to `movmodecc` but for conditional addition. Conditionally move operand 2 or (operands 2 + operand 3) into operand 0 according to the comparison in operand 1. If the comparison is true, operand 2 is moved into operand 0, otherwise (operand 2 + operand 3) is moved.

Store zero or nonzero in operand 0 according to whether a comparison is true. Operand 1 is a comparison operator. Operand 2 and operand 3 are the first and second operand of the comparison, respectively. You specify the mode that operand 0 must have when you write the `match_operand` expression. The compiler automatically sees which mode you have used and supplies an operand of that mode.

The value stored for a true condition must have 1 as its low bit, or else must be negative. Otherwise the instruction is not suitable and you should omit it from the machine description. You describe to the compiler exactly which value is stored by defining the macro `STORE_FLAG_VALUE` (see Section 17.31 [Misc], page 542). If a description cannot be found that can be used for all the possible
comparison operators, you should pick one and use a \texttt{define\_expand} to map all results onto the one you chose.

These operations may \texttt{FAIL}, but should do so only in relatively uncommon cases; if they would \texttt{FAIL} for common cases involving integer comparisons, it is best to restrict the predicates to not allow these operands. Likewise if a given comparison operator will always fail, independent of the operands (for floating-point modes, the \texttt{ordered\_comparison\_operator} predicate is often useful in this case).

If this pattern is omitted, the compiler will generate a conditional branch—for example, it may copy a constant one to the target and branching around an assignment of zero to the target—or a libcall. If the predicate for operand 1 only rejects some operators, it will also try reordering the operands and/or inverting the result value (e.g. by an exclusive OR). These possibilities could be cheaper or equivalent to the instructions used for the \texttt{cstore\_mode4} pattern followed by those required to convert a positive result from \texttt{STORE\_FLAG\_VALUE} to 1; in this case, you can and should make operand 1's predicate reject some operators in the \texttt{cstore\_mode4} pattern, or remove the pattern altogether from the machine description.

\texttt{cbranch\_mode4}

Conditional branch instruction combined with a compare instruction. Operand 0 is a comparison operator. Operand 1 and operand 2 are the first and second operands of the comparison, respectively. Operand 3 is a \texttt{label\_ref} that refers to the label to jump to.

\texttt{jump}

A jump inside a function; an unconditional branch. Operand 0 is the \texttt{label\_ref} of the label to jump to. This pattern name is mandatory on all machines.

\texttt{call}

Subroutine call instruction returning no value. Operand 0 is the function to call; operand 1 is the number of bytes of arguments pushed as a \texttt{const\_int}; operand 2 is the number of registers used as operands.

On most machines, operand 2 is not actually stored into the RTL pattern. It is supplied for the sake of some RISC machines which need to put this information into the assembler code; they can put it in the RTL instead of operand 1.

Operand 0 should be a \texttt{mem RTX} whose address is the address of the function. Note, however, that this address can be a \texttt{symbol\_ref} expression even if it would not be a legitimate memory address on the target machine. If it is also not a valid argument for a call instruction, the pattern for this operation should be a \texttt{define\_expand} (see Section 16.15 [Expander Definitions], page 361) that places the address into a register and uses that register in the call instruction.

\texttt{call\_value}

Subroutine call instruction returning a value. Operand 0 is the hard register in which the value is returned. There are three more operands, the same as the three operands of the \texttt{call} instruction (but with numbers increased by one).

Subroutines that return \texttt{BLKmode} objects use the \texttt{call} insn.
‘call_pop’, ‘call_value_pop’
Similar to ‘call’ and ‘call_value’, except used if defined and if RETURN_POPS_ARGS is nonzero. They should emit a parallel that contains both the function call and a set to indicate the adjustment made to the frame pointer.

For machines where RETURN_POPS_ARGS can be nonzero, the use of these patterns increases the number of functions for which the frame pointer can be eliminated, if desired.

‘untyped_call’
Subroutine call instruction returning a value of any type. Operand 0 is the function to call; operand 1 is a memory location where the result of calling the function is to be stored; operand 2 is a parallel expression where each element is a set expression that indicates the saving of a function return value into the result block.

This instruction pattern should be defined to support __builtin_apply on machines where special instructions are needed to call a subroutine with arbitrary arguments or to save the value returned. This instruction pattern is required on machines that have multiple registers that can hold a return value (i.e. FUNCTION_VALUE_REGNO_P is true for more than one register).

‘return’
Subroutine return instruction. This instruction pattern name should be defined only if a single instruction can do all the work of returning from a function.

Like the ‘mov’ patterns, this pattern is also used after the RTL generation phase. In this case it is to support machines where multiple instructions are usually needed to return from a function, but some class of functions only requires one instruction to implement a return. Normally, the applicable functions are those which do not need to save any registers or allocate stack space.

For such machines, the condition specified in this pattern should only be true when reload_completed is nonzero and the function’s epilogue would only be a single instruction. For machines with register windows, the routine leaf_function_p may be used to determine if a register window push is required.

Machines that have conditional return instructions should define patterns such as

```
(define_insn ""
  (set (pc)
   (if_then_else (match_operator
      0 "comparison_operator"
      [(cc0) (const_int 0)])
   (return)
   (pc)))))
"condition"
"
```

where condition would normally be the same condition specified on the named ‘return’ pattern.

‘untyped_return’
Untyped subroutine return instruction. This instruction pattern should be defined to support __builtin_return on machines where special instructions are needed to return a value of any type.
Operand 0 is a memory location where the result of calling a function with 
`__builtin_apply` is stored; operand 1 is a parallel expression where each 
element is a set expression that indicates the restoring of a function return 
value from the result block.

`nop` No-op instruction. This instruction pattern name should always be defined to 
output a no-op in assembler code. `(const_int 0)` will do as an RTL pattern.

`indirect_jump` An instruction to jump to an address which is operand zero. This pattern name 
is mandatory on all machines.

`casesi` Instruction to jump through a dispatch table, including bounds checking. This 
instruction takes five operands:
1. The index to dispatch on, which has mode SImode.
2. The lower bound for indices in the table, an integer constant.
3. The total range of indices in the table—the largest index minus the smallest 
one (both inclusive).
4. A label that precedes the table itself.
5. A label to jump to if the index has a value outside the bounds.
The table is an `addr_vec` or `addr_diff_vec` inside of a `jump_insn`. The number 
elements in the table is one plus the difference between the upper bound and 
the lower bound.

`tablejump` Instruction to jump to a variable address. This is a low-level capability which 
can be used to implement a dispatch table when there is no `casesi` pattern. 
This pattern requires two operands: the address or offset, and a label which 
should immediately precede the jump table. If the macro `CASE_VECTOR_PC_RELATIVE` 
evaluates to a nonzero value then the first operand is an offset which 
counts from the address of the table; otherwise, it is an absolute address to 
jump to. In either case, the first operand has mode `Pmode`. 
The `tablejump` insn is always the last insn before the jump table it uses. Its 
assembler code normally has no need to use the second operand, but you should 
incorporate it in the RTL pattern so that the jump optimizer will not delete 
the table as unreachable code.

`decrement_and_branch_until_zero` Conditional branch instruction that decrements a register and jumps if the 
register is nonzero. Operand 0 is the register to decrement and test; operand 
1 is the label to jump to if the register is nonzero. See Section 16.13 [Looping 
Patterns], page 358.

This optional instruction pattern is only used by the combiner, typically for 
loops reversed by the loop optimizer when strength reduction is enabled.

`doloop_end` Conditional branch instruction that decrements a register and jumps if the 
register is nonzero. This instruction takes five operands: Operand 0 is the
register to decrement and test; operand 1 is the number of loop iterations as a
const_int or const0_rtx if this cannot be determined until run-time; operand
2 is the actual or estimated maximum number of iterations as a const_int;
operand 3 is the number of enclosed loops as a const_int (an innermost loop
has a value of 1); operand 4 is the label to jump to if the register is nonzero.
See Section 16.13 [Looping Patterns], page 358.
This optional instruction pattern should be defined for machines with low-
overhead looping instructions as the loop optimizer will try to modify suit-
able loops to utilize it. If nested low-overhead looping is not supported, use a
define_expand (see Section 16.15 [Expander Definitions], page 361) and make
the pattern fail if operand 3 is not const1_rtx. Similarly, if the actual or esti-
mated maximum number of iterations is too large for this instruction, make it
fail.

‘doloop_begin’
Companion instruction to doloop_end required for machines that need to per-
form some initialization, such as loading special registers used by a low-overhead
looping instruction. If initialization insns do not always need to be emitted, use a
define_expand (see Section 16.15 [Expander Definitions], page 361) and make
it fail.

‘canonicalize_funcptr_for_compare’
Canonicalize the function pointer in operand 1 and store the result into operand
0.
Operand 0 is always a reg and has mode Pmode; operand 1 may be a reg, mem,
symbol_ref, const_int, etc and also has mode Pmode.
Canonicalization of a function pointer usually involves computing the address
of the function which would be called if the function pointer were used in an
indirect call.
Only define this pattern if function pointers on the target machine can have
different values but still call the same function when used in an indirect call.

‘save_stack_block’
‘save_stack_function’
‘save_stack_nonlocal’
‘restore_stack_block’
‘restore_stack_function’
‘restore_stack_nonlocal’

Most machines save and restore the stack pointer by copying it to or from an
object of mode Pmode. Do not define these patterns on such machines.
Some machines require special handling for stack pointer saves and restores. On
those machines, define the patterns corresponding to the non-standard cases by
using a define_expand (see Section 16.15 [Expander Definitions], page 361)
that produces the required insns. The three types of saves and restores are:

1. ‘save_stack_block’ saves the stack pointer at the start of a block that
allocates a variable-sized object, and ‘restore_stack_block’ restores the
stack pointer when the block is exited.
2. ‘save_stack_function’ and ‘restore_stack_function’ do a similar job for the outermost block of a function and are used when the function allocates variable-sized objects or calls \texttt{alloca}. Only the epilogue uses the restored stack pointer, allowing a simpler save or restore sequence on some machines.

3. ‘save_stack_nonlocal’ is used in functions that contain labels branched to by nested functions. It saves the stack pointer in such a way that the inner function can use ‘restore_stack_nonlocal’ to restore the stack pointer. The compiler generates code to restore the frame and argument pointer registers, but some machines require saving and restoring additional data such as register window information or stack backchains. Place insns in these patterns to save and restore any such required data.

When saving the stack pointer, operand 0 is the save area and operand 1 is the stack pointer. The mode used to allocate the save area defaults to \texttt{Pmode} but you can override that choice by defining the \texttt{STACK_SAVEAREA_MODE} macro (see Section 17.5 [Storage Layout], page 404). You must specify an integral mode, or \texttt{VOIDmode} if no save area is needed for a particular type of save (either because no save is needed or because a machine-specific save area can be used). Operand 0 is the stack pointer and operand 1 is the save area for restore operations. If ‘save_stack_block’ is defined, operand 0 must not be \texttt{VOIDmode} since these saves can be arbitrarily nested.

A save area is a \texttt{mem} that is at a constant offset from \texttt{virtual_stack_vars_rtx} when the stack pointer is saved for use by nonlocal gotos and a \texttt{reg} in the other two cases.

‘allocate_stack’

Subtract (or add if \texttt{STACK_GROWS_DOWNWARD} is undefined) operand 1 from the stack pointer to create space for dynamically allocated data. Store the resultant pointer to this space into operand 0. If you are allocating space from the main stack, do this by emitting a move insn to copy \texttt{virtual_stack_dynamic_rtx} to operand 0. If you are allocating the space elsewhere, generate code to copy the location of the space to operand 0. In the latter case, you must ensure this space gets freed when the corresponding space on the main stack is free.

Do not define this pattern if all that must be done is the subtraction. Some machines require other operations such as stack probes or maintaining the back chain. Define this pattern to emit those operations in addition to updating the stack pointer.

‘check_stack’

If stack checking (see Section 17.10.3 [Stack Checking], page 441) cannot be done on your system by probing the stack, define this pattern to perform the needed check and signal an error if the stack has overflowed. The single operand is the address in the stack farthest from the current stack pointer that you need to validate. Normally, on platforms where this pattern is needed, you would obtain the stack limit from a global or thread-specific variable or register.
'probe_stack'
If stack checking (see Section 17.10.3 [Stack Checking], page 441) can be done on your system by probing the stack but doing it with a “store zero” instruction is not valid or optimal, define this pattern to do the probing differently and signal an error if the stack has overflowed. The single operand is the memory reference in the stack that needs to be probed.

'nonlocal_goto'
Emit code to generate a non-local goto, e.g., a jump from one function to a label in an outer function. This pattern has four arguments, each representing a value to be used in the jump. The first argument is to be loaded into the frame pointer, the second is the address to branch to (code to dispatch to the actual label), the third is the address of a location where the stack is saved, and the last is the address of the label, to be placed in the location for the incoming static chain.

On most machines you need not define this pattern, since GCC will already generate the correct code, which is to load the frame pointer and static chain, restore the stack (using the 'restore_stack_nonlocal' pattern, if defined), and jump indirectly to the dispatcher. You need only define this pattern if this code will not work on your machine.

'nonlocal_goto_receiver'
This pattern, if defined, contains code needed at the target of a nonlocal goto after the code already generated by GCC. You will not normally need to define this pattern. A typical reason why you might need this pattern is if some value, such as a pointer to a global table, must be restored when the frame pointer is restored. Note that a nonlocal goto only occurs within a unit-of-translation, so a global table pointer that is shared by all functions of a given module need not be restored. There are no arguments.

'exception_receiver'
This pattern, if defined, contains code needed at the site of an exception handler that isn’t needed at the site of a nonlocal goto. You will not normally need to define this pattern. A typical reason why you might need this pattern is if some value, such as a pointer to a global table, must be restored after control flow is branched to the handler of an exception. There are no arguments.

'builtin_setjmp_setup'
This pattern, if defined, contains additional code needed to initialize the jmp_buf. You will not normally need to define this pattern. A typical reason why you might need this pattern is if some value, such as a pointer to a global table, must be restored. Though it is preferred that the pointer value be recalculated if possible (given the address of a label for instance). The single argument is a pointer to the jmp_buf. Note that the buffer is five words long and that the first three are normally used by the generic mechanism.

'builtin_setjmp_receiver'
This pattern, if defined, contains code needed at the site of a built-in setjmp that isn’t needed at the site of a nonlocal goto. You will not normally need to define this pattern. A typical reason why you might need this pattern is if
some value, such as a pointer to a global table, must be restored. It takes one argument, which is the label to which builtin_longjmp transferred control; this pattern may be emitted at a small offset from that label.

`builtin_longjmp`

This pattern, if defined, performs the entire action of the longjmp. You will not normally need to define this pattern unless you also define builtin_setjmp_setup. The single argument is a pointer to the jmp_buf.

`eh_return`

This pattern, if defined, affects the way __builtin_eh_return, and thence the call frame exception handling library routines, are built. It is intended to handle non-trivial actions needed along the abnormal return path.

The address of the exception handler to which the function should return is passed as operand to this pattern. It will normally need to copied by the pattern to some special register or memory location. If the pattern needs to determine the location of the target call frame in order to do so, it may use EH_RETURN_STACKADJ_RTX, if defined; it will have already been assigned.

If this pattern is not defined, the default action will be to simply copy the return address to EH_RETURN_HANDLER_RTX. Either that macro or this pattern needs to be defined if call frame exception handling is to be used.

`prologue`

This pattern, if defined, emits RTL for entry to a function. The function entry is responsible for setting up the stack frame, initializing the frame pointer register, saving callee saved registers, etc.

Using a prologue pattern is generally preferred over defining TARGET_ASM_FUNCTION_PROLOGUE to emit assembly code for the prologue.

The prologue pattern is particularly useful for targets which perform instruction scheduling.

`epilogue`

This pattern emits RTL for exit from a function. The function exit is responsible for deallocating the stack frame, restoring callee saved registers and emitting the return instruction.

Using an epilogue pattern is generally preferred over defining TARGET_ASM_FUNCTION_EPILOGUE to emit assembly code for the epilogue.

The epilogue pattern is particularly useful for targets which perform instruction scheduling or which have delay slots for their return instruction.

`sibcall_epilogue`

This pattern, if defined, emits RTL for exit from a function without the final branch back to the calling function. This pattern will be emitted before any sibling call (aka tail call) sites.

The sibcall_epilogue pattern must not clobber any arguments used for parameter passing or any stack slots for arguments passed to the current function.

`trap`

This pattern, if defined, signals an error, typically by causing some kind of signal to be raised. Among other places, it is used by the Java front end to signal ‘invalid array index’ exceptions.
Conditional trap instruction. Operand 0 is a piece of RTL which performs a comparison, and operands 1 and 2 are the arms of the comparison. Operand 3 is the trap code, an integer.

A typical `ctrap` pattern looks like

```
(define_insn "ctraşi4"
  [(trap_if (match_operator 0 "trap_operator"
                             [(match_operand 1 "register_operand")
                             (match_operand 2 "immediate_operand")
                             (match_operand 3 "const_int_operand" "i"))
    ""
    "...")]
```

This pattern, if defined, emits code for a non-faulting data prefetch instruction. Operand 0 is the address of the memory to prefetch. Operand 1 is a constant 1 if the prefetch is preparing for a write to the memory address, or a constant 0 otherwise. Operand 2 is the expected degree of temporal locality of the data and is a value between 0 and 3, inclusive; 0 means that the data has no temporal locality, so it need not be left in the cache after the access; 3 means that the data has a high degree of temporal locality and should be left in all levels of cache possible; 1 and 2 mean, respectively, a low or moderate degree of temporal locality.

Targets that do not support write prefetches or locality hints can ignore the values of operands 1 and 2.

This pattern defines a pseudo insn that prevents the instruction scheduler from moving instructions across the boundary defined by the blockage insn. Normally an UNSPEC_VOLATILE pattern.

If the target memory model is not fully synchronous, then this pattern should be defined to an instruction that orders both loads and stores before the instruction with respect to loads and stores after the instruction. This pattern has no operands.

This pattern, if defined, emits code for an atomic compare-and-swap operation. Operand 1 is the memory on which the atomic operation is performed. Operand 2 is the “old” value to be compared against the current contents of the memory location. Operand 3 is the “new” value to store in the memory if the compare succeeds. Operand 0 is the result of the operation; it should contain the contents of the memory before the operation. If the compare succeeds, this should obviously be a copy of operand 2.

This pattern must show that both operand 0 and operand 1 are modified.

This pattern must issue any memory barrier instructions such that all memory operations before the atomic operation occur before the atomic operation and all memory operations after the atomic operation occur after the atomic operation.
For targets where the success or failure of the compare-and-swap operation is available via the status flags, it is possible to avoid a separate compare operation and issue the subsequent branch or store-flag operation immediately after the compare-and-swap. To this end, GCC will look for a \texttt{MODE\_CC} set in the output of \texttt{sync\_compare\_and\_swapmode}; if the machine description includes such a set, the target should also define special \texttt{cbranchcc4} and/or \texttt{cstorecc4} instructions. GCC will then be able to take the destination of the \texttt{MODE\_CC} set and pass it to the \texttt{cbranchcc4} or \texttt{cstorecc4} pattern as the first operand of the comparison (the second will be \texttt{(const\_int 0)}).

\texttt{'sync\_addmode'}, \texttt{'sync\_submode'}
\texttt{'sync\_iormode'}, \texttt{'sync\_andmode'}
\texttt{'sync\_xormode'}, \texttt{'sync\_nandmode'}

These patterns emit code for an atomic operation on memory. Operand 0 is the memory on which the atomic operation is performed. Operand 1 is the second operand to the binary operator.

This pattern must issue any memory barrier instructions such that all memory operations before the atomic operation occur before the atomic operation and all memory operations after the atomic operation occur after the atomic operation. If these patterns are not defined, the operation will be constructed from a compare-and-swap operation, if defined.

\texttt{'sync\_old\_addmode'}, \texttt{'sync\_old\_submode'}
\texttt{'sync\_old\_iormode'}, \texttt{'sync\_old\_andmode'}
\texttt{'sync\_old\_xormode'}, \texttt{'sync\_old\_nandmode'}

These patterns are emit code for an atomic operation on memory, and return the value that the memory contained before the operation. Operand 0 is the result value, operand 1 is the memory on which the atomic operation is performed, and operand 2 is the second operand to the binary operator.

This pattern must issue any memory barrier instructions such that all memory operations before the atomic operation occur before the atomic operation and all memory operations after the atomic operation occur after the atomic operation. If these patterns are not defined, the operation will be constructed from a compare-and-swap operation, if defined.

\texttt{'sync\_new\_addmode'}, \texttt{'sync\_new\_submode'}
\texttt{'sync\_new\_iormode'}, \texttt{'sync\_new\_andmode'}
\texttt{'sync\_new\_xormode'}, \texttt{'sync\_new\_nandmode'}

These patterns are like their \texttt{sync\_old\_op} counterparts, except that they return the value that exists in the memory location after the operation, rather than before the operation.

\texttt{'sync\_lock\_test\_and\_setmode'}

This pattern takes two forms, based on the capabilities of the target. In either case, operand 0 is the result of the operand, operand 1 is the memory on which the atomic operation is performed, and operand 2 is the value to set in the lock.

In the ideal case, this operation is an atomic exchange operation, in which the previous value in memory operand is copied into the result operand, and the value operand is stored in the memory operand.
For less capable targets, any value operand that is not the constant 1 should be rejected with FAIL. In this case the target may use an atomic test-and-set bit operation. The result operand should contain 1 if the bit was previously set and 0 if the bit was previously clear. The true contents of the memory operand are implementation defined.

This pattern must issue any memory barrier instructions such that the pattern as a whole acts as an acquire barrier, that is all memory operations after the pattern do not occur until the lock is acquired.

If this pattern is not defined, the operation will be constructed from a compare-and-swap operation, if defined.

'sync_lock_release_mode'
This pattern, if defined, releases a lock set by sync_lock_test_and_setmode. Operand 0 is the memory that contains the lock; operand 1 is the value to store in the lock.

If the target doesn’t implement full semantics for sync_lock_test_and_setmode, any value operand which is not the constant 0 should be rejected with FAIL, and the true contents of the memory operand are implementation defined.

This pattern must issue any memory barrier instructions such that the pattern as a whole acts as a release barrier, that is the lock is released only after all previous memory operations have completed.

If this pattern is not defined, then a memory_barrier pattern will be emitted, followed by a store of the value to the memory operand.

'stack_protect_set'
This pattern, if defined, moves a ptr_mode value from the memory in operand 1 to the memory in operand 0 without leaving the value in a register afterward. This is to avoid leaking the value some place that an attacker might use to rewrite the stack guard slot after having clobbered it.

If this pattern is not defined, then a plain move pattern is generated.

'stack_protect_test'
This pattern, if defined, compares a ptr_mode value from the memory in operand 1 with the memory in operand 0 without leaving the value in a register afterward and branches to operand 2 if the values weren’t equal.

If this pattern is not defined, then a plain compare pattern and conditional branch pattern is used.

'clear_cache'
This pattern, if defined, flushes the instruction cache for a region of memory. The region is bounded to by the Pmode pointers in operand 0 inclusive and operand 1 exclusive.

If this pattern is not defined, a call to the library function __clear_cache is used.
16.10 When the Order of Patterns Matters

Sometimes an insn can match more than one instruction pattern. Then the pattern that appears first in the machine description is the one used. Therefore, more specific patterns (patterns that will match fewer things) and faster instructions (those that will produce better code when they do match) should usually go first in the description.

In some cases the effect of ordering the patterns can be used to hide a pattern when it is not valid. For example, the 68000 has an instruction for converting a fullword to floating point and another for converting a byte to floating point. An instruction converting an integer to floating point could match either one. We put the pattern to convert the fullword first to make sure that one will be used rather than the other. (Otherwise a large integer might be generated as a single-byte immediate quantity, which would not work.) Instead of using this pattern ordering it would be possible to make the pattern for convert-a-byte smart enough to deal properly with any constant value.

16.11 Interdependence of Patterns

In some cases machines support instructions identical except for the machine mode of one or more operands. For example, there may be “sign-extend halfword” and “sign-extend byte” instructions whose patterns are

\[
\begin{align*}
\text{(set (match_operand:SI 0 \ldots))} \\
\text{(extend:SI (match_operand:HI 1 \ldots)))}
\end{align*}
\]

\[
\begin{align*}
\text{(set (match_operand:SI 0 \ldots))} \\
\text{(extend:SI (match_operand:QI 1 \ldots)))}
\end{align*}
\]

Constant integers do not specify a machine mode, so an instruction to extend a constant value could match either pattern. The pattern it actually will match is the one that appears first in the file. For correct results, this must be the one for the widest possible mode (\texttt{HImode}, here). If the pattern matches the \texttt{QImode} instruction, the results will be incorrect if the constant value does not actually fit that mode.

Such instructions to extend constants are rarely generated because they are optimized away, but they do occasionally happen in nonoptimized compilations.

If a constraint in a pattern allows a constant, the reload pass may replace a register with a constant permitted by the constraint in some cases. Similarly for memory references. Because of this substitution, you should not provide separate patterns for increment and decrement instructions. Instead, they should be generated from the same pattern that supports register-register add insns by examining the operands and generating the appropriate machine instruction.

16.12 Defining Jump Instruction Patterns

GCC does not assume anything about how the machine realizes jumps. The machine description should define a single pattern, usually a \texttt{define\_expand}, which expands to all the required insns.

Usually, this would be a comparison insn to set the condition code and a separate branch insn testing the condition code and branching or not according to its value. For many machines, however, separating compares and branches is limiting, which is why the more
A flexible approach with one `define_expand` is used in GCC. The machine description becomes clearer for architectures that have compare-and-branch instructions but no condition code. It also works better when different sets of comparison operators are supported by different kinds of conditional branches (e.g. integer vs. floating-point), or by conditional branches with respect to conditional stores.

Two separate insns are always used if the machine description represents a condition code register using the legacy RTL expression `cc0`, and on most machines that use a separate condition code register (see Section 17.16 [Condition Code], page 476). For machines that use `(cc0)`, in fact, the set and use of the condition code must be separate and adjacent\(^1\), thus allowing flags in `cc_status` to be used (see Section 17.16 [Condition Code], page 476) and so that the comparison and branch insns could be located from each other by using the functions `prev_cc0_setter` and `next_cc0_user`.

Even in this case having a single entry point for conditional branches is advantageous, because it handles equally well the case where a single comparison instruction records the results of both signed and unsigned comparison of the given operands (with the branch insns coming in distinct signed and unsigned flavors) as in the x86 or SPARC, and the case where there are distinct signed and unsigned compare instructions and only one set of conditional branch instructions as in the PowerPC.

### 16.13 Defining Looping Instruction Patterns

Some machines have special jump instructions that can be utilized to make loops more efficient. A common example is the 68000 ‘`dbra`’ instruction which performs a decrement of a register and a branch if the result was greater than zero. Other machines, in particular digital signal processors (DSPs), have special block repeat instructions to provide low-overhead loop support. For example, the TI TMS320C3x/C4x DSPs have a block repeat instruction that loads special registers to mark the top and end of a loop and to count the number of loop iterations. This avoids the need for fetching and executing a `dbra`-like instruction and avoids pipeline stalls associated with the jump.

GCC has three special named patterns to support low overhead looping. They are `decrement_and_branch_until_zero`, `doloop_begin`, and `doloop_end`. The first pattern, `decrement_and_branch_until_zero`, is not emitted during RTL generation but may be emitted during the instruction combination phase. This requires the assistance of the loop optimizer, using information collected during strength reduction, to reverse a loop to count down to zero. Some targets also require the loop optimizer to add a `REG_NONNEG` note to indicate that the iteration count is always positive. This is needed if the target performs a signed loop termination test. For example, the 68000 uses a pattern similar to the following for its `dbra` instruction:

\(^1\) *Note* insns can separate them, though.
Chapter 16: Machine Descriptions

(define_insn "decrement_and_branch_until_zero"
 [(set (pc)
   (if_then_else
    (ge (plus:SI (match_operand:SI 0 "general_operand" "+d*am")
         (const_int -1))
    (const_int 0))
    (label_ref (match_operand 1 "" ")
        (pc))]
   (set (match_dup 0)
        (plus:SI (match_dup 0)
                  (const_int -1))))]
 "find_reg_note (insn, REG_NONNEG, 0)"
 "..."

Note that since the insn is both a jump insn and has an output, it must deal with its own reloads, hence the ‘m’ constraints. Also note that since this insn is generated by the instruction combination phase combining two sequential insns together into an implicit parallel insn, the iteration counter needs to be biased by the same amount as the decrement operation, in this case −1. Note that the following similar pattern will not be matched by the combiner.

(define_insn "decrement_and_branch_until_zero"
 [(set (pc)
   (if_then_else
    (ge (match_operand:SI 0 "general_operand" "+d*am")
         (const_int 1))
    (label_ref (match_operand 1 "" ")
        (pc))]
   (set (match_dup 0)
        (plus:SI (match_dup 0)
                  (const_int -1))))]
 "find_reg_note (insn, REG_NONNEG, 0)"
 "...

The other two special looping patterns, ‘doloop_begin’ and ‘doloop_end’, are emitted by the loop optimizer for certain well-behaved loops with a finite number of loop iterations using information collected during strength reduction.

The ‘doloop_end’ pattern describes the actual looping instruction (or the implicit looping operation) and the ‘doloop_begin’ pattern is an optional companion pattern that can be used for initialization needed for some low-overhead looping instructions.

Note that some machines require the actual looping instruction to be emitted at the top of the loop (e.g., the TMS320C3x/C4x DSPs). Emitting the true RTL for a looping instruction at the top of the loop can cause problems with flow analysis. So instead, a dummy doloop insn is emitted at the end of the loop. The machine dependent reorg pass checks for the presence of this doloop insn and then searches back to the top of the loop, where it inserts the true looping insn (provided there are no instructions in the loop which would cause problems). Any additional labels can be emitted at this point. In addition, if the desired special iteration counter register was not allocated, this machine dependent reorg pass could emit a traditional compare and jump instruction pair.

The essential difference between the ‘decrement_and_branch_until_zero’ and the ‘doloop_end’ patterns is that the loop optimizer allocates an additional pseudo register for the latter as an iteration counter. This pseudo register cannot be used within the loop (i.e., general induction variables cannot be derived from it), however, in many cases the loop induction variable may become redundant and removed by the flow pass.
16.14 Canonicalization of Instructions

There are often cases where multiple RTL expressions could represent an operation performed by a single machine instruction. This situation is most commonly encountered with logical, branch, and multiply-accumulate instructions. In such cases, the compiler attempts to convert these multiple RTL expressions into a single canonical form to reduce the number of insn patterns required.

In addition to algebraic simplifications, following canonicalizations are performed:

- For commutative and comparison operators, a constant is always made the second operand. If a machine only supports a constant as the second operand, only patterns that match a constant in the second operand need be supplied.
- For associative operators, a sequence of operators will always chain to the left; for instance, only the left operand of an integer plus can itself be a plus. and, ior, xor, plus, mult, smin, smax, umin, and umax are associative when applied to integers, and sometimes to floating-point.
- For these operators, if only one operand is a neg, not, mult, plus, or minus expression, it will be the first operand.
- In combinations of neg, mult, plus, and minus, the neg operations (if any) will be moved inside the operations as far as possible. For instance, (neg (mult A B)) is canonicalized as (mult (neg A) B), but (plus (mult (neg B) C) A) is canonicalized as (minus A (mult B C)).
- For the compare operator, a constant is always the second operand if the first argument is a condition code register or (cc0).
- An operand of neg, not, mult, plus, or minus is made the first operand under the same conditions as above.
- (ltu (plus a b) b) is converted to (ltu (plus a b) a). Likewise with geu instead of ltu.
- (minus x (const_int n)) is converted to (plus x (const_int -n)).
- Within address computations (i.e., inside mem), a left shift is converted into the appropriate multiplication by a power of two.
- De Morgan’s Law is used to move bitwise negation inside a bitwise logical-and or logical-or operation. If this results in only one operand being a not expression, it will be the first one.

A machine that has an instruction that performs a bitwise logical-and of one operand with the bitwise negation of the other should specify the pattern for that instruction as

```c
(define_insn ""
  [(set (match_operand:m 0 ...)  
    (and:m (not:m (match_operand:m 1 ...))  
      (match_operand:m 2 ...)))]
  
  "...
  "..."
)
```

Similarly, a pattern for a “NAND” instruction should be written

```c
(define_insn ""
  [(set (match_operand:m 0 ...)  
    (ior:m (not:m (match_operand:m 1 ...))  
```

```c
```

```c
```
In both cases, it is not necessary to include patterns for the many logically equivalent RTL expressions.

- The only possible RTL expressions involving both bitwise exclusive-or and bitwise negation are \((\text{xor:} m \ x \ y)\) and \((\text{not:} m \ (\text{xor:} m \ x \ y))\).

- The sum of three items, one of which is a constant, will only appear in the form \((\text{plus:} m \ (\text{plus:} m \ x \ y) \ \text{constant})\).

- Equality comparisons of a group of bits (usually a single bit) with zero will be written using \text{zero_extract} rather than the equivalent \text{and} or \text{sign_extract} operations.

Further canonicalization rules are defined in the function \text{commutative_operand_precedence} in 'gcc/rtlanal.c'.

### 16.15 Defining RTL Sequences for Code Generation

On some target machines, some standard pattern names for RTL generation cannot be handled with single insn, but a sequence of RTL insns can represent them. For these target machines, you can write a \text{define_expand} to specify how to generate the sequence of RTL.

A \text{define_expand} is an RTL expression that looks almost like a \text{define_insn}; but, unlike the latter, a \text{define_expand} is used only for RTL generation and it can produce more than one RTL insn.

A \text{define_expand} RTX has four operands:

- The name. Each \text{define_expand} must have a name, since the only use for it is to refer to it by name.

- The RTL template. This is a vector of RTL expressions representing a sequence of separate instructions. Unlike \text{define_insn}, there is no implicit surrounding \text{PARALLEL}.

- The condition, a string containing a C expression. This expression is used to express how the availability of this pattern depends on subclasses of target machine, selected by command-line options when GCC is run. This is just like the condition of a \text{define_insn} that has a standard name. Therefore, the condition (if present) may not depend on the data in the insn being matched, but only the target-machine-type flags. The compiler needs to test these conditions during initialization in order to learn exactly which named instructions are available in a particular run.

- The preparation statements, a string containing zero or more C statements which are to be executed before RTL code is generated from the RTL template.

Usually these statements prepare temporary registers for use as internal operands in the RTL template, but they can also generate RTL insns directly by calling routines such as \text{emit_insn}, etc. Any such insns precede the ones that come from the RTL template.

Every RTL insn emitted by a \text{define_expand} must match some \text{define_insn} in the machine description. Otherwise, the compiler will crash when trying to generate code for the insn or trying to optimize it.
The RTL template, in addition to controlling generation of RTL insns, also describes the operands that need to be specified when this pattern is used. In particular, it gives a predicate for each operand.

A true operand, which needs to be specified in order to generate RTL from the pattern, should be described with a `match_operand` in its first occurrence in the RTL template. This enters information on the operand's predicate into the tables that record such things. GCC uses the information to preload the operand into a register if that is required for valid RTL code. If the operand is referred to more than once, subsequent references should use `match_dup`.

The RTL template may also refer to internal “operands” which are temporary registers or labels used only within the sequence made by the `define_expand`. Internal operands are substituted into the RTL template with `match_dup`, never with `match_operand`. The values of the internal operands are not passed in as arguments by the compiler when it requests use of this pattern. Instead, they are computed within the pattern, in the preparation statements. These statements compute the values and store them into the appropriate elements of `operands` so that `match_dup` can find them.

There are two special macros defined for use in the preparation statements: `DONE` and `FAIL`. Use them with a following semicolon, as a statement.

**DONE**  
Use the `DONE` macro to end RTL generation for the pattern. The only RTL insns resulting from the pattern on this occasion will be those already emitted by explicit calls to `emit_insn` within the preparation statements; the RTL template will not be generated.

**FAIL**  
Make the pattern fail on this occasion. When a pattern fails, it means that the pattern was not truly available. The calling routines in the compiler will try other strategies for code generation using other patterns.

Failure is currently supported only for binary (addition, multiplication, shifting, etc.) and bit-field (`extv`, `extzv`, and `insv`) operations.

If the preparation falls through (invokes neither `DONE` nor `FAIL`), then the `define_expand` acts like a `define_insn` in that the RTL template is used to generate the insn.

The RTL template is not used for matching, only for generating the initial insn list. If the preparation statement always invokes `DONE` or `FAIL`, the RTL template may be reduced to a simple list of operands, such as this example:

```
(define_expand "addsi3"
  [(match_operand:SI 0 "register_operand" ")
   (match_operand:SI 1 "register_operand" ")
   (match_operand:SI 2 "register_operand" "]
     "

   { 
       handle_add (operands[0], operands[1], operands[2]);
       DONE;
   })"
```

Here is an example, the definition of left-shift for the SPUR chip:

```
(define_expand "ashlsi3"
  [(set (match_operand:SI 0 "register_operand" ")
       (ashift:SI
```
This example uses `define_expand` so that it can generate an RTL insn for shifting when the shift-count is in the supported range of 0 to 3 but fail in other cases where machine insns aren’t available. When it fails, the compiler tries another strategy using different patterns (such as, a library call).

If the compiler were able to handle nontrivial condition-strings in patterns with names, then it would be possible to use a `define_insn` in that case. Here is another case (zero-extension on the 68000) which makes more use of the power of `defineexpand`:

```asm
(define_expand "zero_extendhisi2"
  [(set (match_operand:SI 0 "general_operand" "")
        (const_int 0))
   (set (strict_low_part
        (subreg:HI
         (match_dup 0)
        0))

   (match_operand:HI 1 "general_operand" "")])

"operands[1] = make_safe_from (operands[1], operands[0]);"
```

Here two RTL insns are generated, one to clear the entire output operand and the other to copy the input operand into its low half. This sequence is incorrect if the input operand refers to [the old value of] the output operand, so the preparation statement makes sure this isn’t so. The function `make_safe_from` copies the `operands[1]` into a temporary register if it refers to `operands[0]`. It does this by emitting another RTL insn.

Finally, a third example shows the use of an internal operand. Zero-extension on the SPUR chip is done by `and-ing` the result against a halfword mask. But this mask cannot be represented by a `const_int` because the constant value is too large to be legitimate on this machine. So it must be copied into a register with `force_reg` and then the register used in the `and`.

```asm
(define_expand "zero_extendhisi2"
  [(set (match_operand:SI 0 "register_operand" "")
        (and:SI (subreg:SI
                 (match_operand:HI 1 "register_operand" "")
        0)
   (match_dup 2)))]

"operands[2] = force_reg (SImode, GEN_INT (65535)); ")
```

**Note:** If the `define_expand` is used to serve a standard binary or unary arithmetic operation or a bit-field operation, then the last insn it generates must not be a `code_label`, `barrier` or `note`. It must be an `insn`, `jump_insn` or `call_insn`. If you don’t need a real insn at the end, emit an insn to copy the result of the operation into itself. Such an insn will generate no code, but it can avoid problems in the compiler.
16.16 Defining How to Split Instructions

There are two cases where you should specify how to split a pattern into multiple insns. On machines that have instructions requiring delay slots (see Section 16.19.7 [Delay Slots], page 379) or that have instructions whose output is not available for multiple cycles (see Section 16.19.8 [Processor pipeline description], page 380), the compiler phases that optimize these cases need to be able to move insns into one-instruction delay slots. However, some insns may generate more than one machine instruction. These insns cannot be placed into a delay slot.

Often you can rewrite the single insn as a list of individual insns, each corresponding to one machine instruction. The disadvantage of doing so is that it will cause the compilation to be slower and require more space. If the resulting insns are too complex, it may also suppress some optimizations. The compiler splits the insn if there is a reason to believe that it might improve instruction or delay slot scheduling.

The insn combiner phase also splits putative insns. If three insns are merged into one insn with a complex expression that cannot be matched by some `define_insn` pattern, the combiner phase attempts to split the complex pattern into two insns that are recognized. Usually it can break the complex pattern into two patterns by splitting out some subexpression. However, in some other cases, such as performing an addition of a large constant in two insns on a RISC machine, the way to split the addition into two insns is machine-dependent.

The `define_split` definition tells the compiler how to split a complex insn into several simpler insns. It looks like this:

```
(define_split
  [insn-pattern]
  "condition"
  [new_insn-pattern-1
    new_insn-pattern-2
    ...
  ]
  "preparation-statements")
```

`insn-pattern` is a pattern that needs to be split and `condition` is the final condition to be tested, as in a `define_insn`. When an insn matching `insn-pattern` and satisfying `condition` is found, it is replaced in the insn list with the insns given by `new_insn-pattern-1`, `new_insn-pattern-2`, etc.

The `preparation-statements` are similar to those statements that are specified for `define_expand` (see Section 16.15 [Expander Definitions], page 361) and are executed before the new RTL is generated to prepare for the generated code or emit some insns whose pattern is not fixed. Unlike those in `define_expand`, however, these statements must not generate any new pseudo-registers. Once reload has completed, they also must not allocate any space in the stack frame.

Patterns are matched against `insn-pattern` in two different circumstances. If an insn needs to be split for delay slot scheduling or insn scheduling, the insn is already known to be valid, which means that it must have been matched by some `define_insn` and, if `reload_completed` is nonzero, is known to satisfy the constraints of that `define_insn`. In that case, the new insn patterns must also be insns that are matched by some `define_insn` and, if `reload_completed` is nonzero, must also satisfy the constraints of those definitions.
As an example of this usage of define_split, consider the following example from `a29k.md`, which splits a sign_extend from HImode to SImode into a pair of shift insns:

```lisp
(define_split
  [(set (match_operand:SI 0 "gen_reg_operand" "")
       (sign_extend:SI (match_operand:HI 1 "gen_reg_operand" "")))
   ""
  [(set (match_dup 0)
         (ashift:SI (match_dup 1)
                    (const_int 16)))
   (set (match_dup 0)
         (ashiftrt:SI (match_dup 0)
                      (const_int 16)))
   "
   { operands[1] = gen_lowpart (SImode, operands[1]); }"
]
```

When the combiner phase tries to split an insn pattern, it is always the case that the pattern is not matched by any define_insn. The combiner pass first tries to split a single set expression and then the same set expression inside a parallel, but followed by a clobber of a pseudo-reg to use as a scratch register. In these cases, the combiner expects exactly two new insn patterns to be generated. It will verify that these patterns match some define_insn definitions, so you need not do this test in the define_split (of course, there is no point in writing a define_split that will never produce insns that match).

Here is an example of this use of define_split, taken from `rs6000.md`:

```lisp
(define_split
  [(set (match_operand:SI 0 "gen_reg_operand" "")
       (plus:SI (match_operand:SI 1 "gen_reg_operand" ")
                (match_operand:SI 2 "non_add_cint_operand" "")))
   ""
  [(set (match_dup 0) (plus:SI (match_dup 1) (match_dup 3)))
   (set (match_dup 0) (plus:SI (match_dup 0) (match_dup 4)))
   "
   { int low = INTVAL (operands[2]) & 0xffff;
     int high = (unsigned) INTVAL (operands[2]) >> 16;

     if (low & 0x8000)
       high++, low |= 0xffff0000;

     operands[3] = GEN_INT (high << 16);
     operands[4] = GEN_INT (low);
   }"
]
```

Here the predicate non_add_cint_operand matches any const_int that is not a valid operand of a single add insn. The add with the smaller displacement is written so that it can be substituted into the address of a subsequent operation.

An example that uses a scratch register, from the same file, generates an equality comparison of a register and a large constant:

```lisp
(define_split
  [(set (match_operand:CC 0 "cc_reg_operand" "")
       (compare:CC (match_operand:SI 1 "gen_reg_operand" ")
                    (match_operand:SI 2 "non_short_cint_operand" "")))
   "find_single_use (operands[0], insn, 0)
   && (GET_CODE (*find_single_use (operands[0], insn, 0)) == EQ
        || GET_CODE (*find_single_use (operands[0], insn, 0)) == NE)"
```
```c
(set (match_dup 3) (xor:SI (match_dup 1) (match_dup 4)))
(set (match_dup 0) (compare:CC (match_dup 3) (match_dup 5))))

/* Get the constant we are comparing against, C, and see what it
 looks like sign-extended to 16 bits. Then see what constant
 could be XOR'ed with C to get the sign-extended value. */

int c = INTVAL (operands[2]);
int sextc = (c << 16) >> 16;
int xorv = c ^ sextc;

operands[4] = GEN_INT (xorv);
operands[5] = GEN_INT (sextc);
}
```

To avoid confusion, don’t write a single `define_split` that accepts some insns that match some `define_insn` as well as some insns that don’t. Instead, write two separate `define_split` definitions, one for the insns that are valid and one for the insns that are not valid.

The splitter is allowed to split jump instructions into sequence of jumps or create new jumps in while splitting non-jump instructions. As the central flowgraph and branch prediction information needs to be updated, several restriction apply.

Splitting of jump instruction into sequence that over by another jump instruction is always valid, as compiler expect identical behavior of new jump. When new sequence contains multiple jump instructions or new labels, more assistance is needed. Splitter is required to create only unconditional jumps, or simple conditional jump instructions. Additionally it must attach a `REG_BR_PROB` note to each conditional jump. A global variable `split_branch_probability` holds the probability of the original branch in case it was a simple conditional jump, −1 otherwise. To simplify recomputing of edge frequencies, the new sequence is required to have only forward jumps to the newly created labels.

For the common case where the pattern of a `define_split` exactly matches the pattern of a `define_insn`, use `define_insn_and_split`. It looks like this:

```c
(define_insn_and_split
 [insn-pattern]  
 "condition"
 [output-template]
 "split-condition"
 [new_insn-pattern-1
     new_insn-pattern-2
     ...
     ]
 "preparation-statements"
 [insn-attributes])
```

`insn-pattern`, `condition`, `output-template`, and `insn-attributes` are used as in `define_insn`. The `new_insn-pattern` vector and the `preparation-statements` are used as in a `define_split`. The `split-condition` is also used as in `define_split`, with the additional behavior that if the condition starts with ‘&&’, the condition used for the split will be the constructed as a logical “and” of the split condition with the insn condition. For example, from i386.md:

```c
(define_insn_and_split "zero_extendhisi2_and"
 [(set (match_operand:SI 0 "register_operand" "=r")
     (zero_extend:SI (match_operand:HI 1 "register_operand" "0")))))
```
In this case, the actual split condition will be `TARGET_ZERO_EXTEND_WITH_AND && !optimize_size`.

The `define_insn_and_split` construction provides exactly the same functionality as two separate `define_insn` and `define_split` patterns. It exists for compactness, and as a maintenance tool to prevent having to ensure the two patterns’ templates match.

### 16.17 Including Patterns in Machine Descriptions.

The `include` pattern tells the compiler tools where to look for patterns that are in files other than in the file `.md`. This is used only at build time and there is no preprocessing allowed.

It looks like:

```c
(include pathname)
```

For example:

```c
(include "filestuff")
```

Where `pathname` is a string that specifies the location of the file, specifies the include file to be in `gcc/config/target/filestuff`. The directory `gcc/config/target` is regarded as the default directory.

Machine descriptions may be split up into smaller more manageable subsections and placed into subdirectories.

By specifying:

```c
(include "BOGUS/filestuff")
```

the include file is specified to be in `gcc/config/target/BOGUS/filestuff`.

Specifying an absolute path for the include file such as:

```c
(include "/u2/BOGUS/filestuff")
```

is permitted but is not encouraged.

### 16.17.1 RTL Generation Tool Options for Directory Search

The `--I dir` option specifies directories to search for machine descriptions. For example:
Add the directory `dir` to the head of the list of directories to be searched for header files. This can be used to override a system machine definition file, substituting your own version, since these directories are searched before the default machine description file directories. If you use more than one `-I` option, the directories are scanned in left-to-right order; the standard default directory come after.

### 16.18 Machine-Specific Peephole Optimizers

In addition to instruction patterns the `md` file may contain definitions of machine-specific peephole optimizations.

The combiner does not notice certain peephole optimizations when the data flow in the program does not suggest that it should try them. For example, sometimes two consecutive insns related in purpose can be combined even though the second one does not appear to use a register computed in the first one. A machine-specific peephole optimizer can detect such opportunities.

There are two forms of peephole definitions that may be used. The original `define_peephole` is run at assembly output time to match insns and substitute assembly text. Use of `define_peephole` is deprecated.

A newer `define_peephole2` matches insns and substitutes new insns. The `peephole2` pass is run after register allocation but before scheduling, which may result in much better code for targets that do scheduling.

#### 16.18.1 RTL to Text Peephole Optimizers

A definition looks like this:

```plaintext
(define_peephole
  [insn-pattern-1
   insn-pattern-2
   ...]
  "condition"
  "template"
  "optional-insn-attributes")
```

The last string operand may be omitted if you are not using any machine-specific information in this machine description. If present, it must obey the same rules as in a `define_insn`.

In this skeleton, `insn-pattern-1` and so on are patterns to match consecutive insns. The optimization applies to a sequence of insns when `insn-pattern-1` matches the first one, `insn-pattern-2` matches the next, and so on.

Each of the insns matched by a peephole must also match a `define_insn`. Peepholes are checked only at the last stage just before code generation, and only optionally. Therefore, any insn which would match a peephole but no `define_insn` will cause a crash in code generation in an unoptimized compilation, or at various optimization stages.

The operands of the insns are matched with `match_operands`, `match_operator`, and `match_dup`, as usual. What is not usual is that the operand numbers apply to all the insn patterns in the definition. So, you can check for identical operands in two insns by using `match_operand` in one insn and `match_dup` in the other.
The operand constraints used in `match_operand` patterns do not have any direct effect on the applicability of the peephole, but they will be validated afterward, so make sure your constraints are general enough to apply whenever the peephole matches. If the peephole matches but the constraints are not satisfied, the compiler will crash.

It is safe to omit constraints in all the operands of the peephole; or you can write constraints which serve as a double-check on the criteria previously tested.

Once a sequence of insns matches the patterns, the `condition` is checked. This is a C expression which makes the final decision whether to perform the optimization (we do so if the expression is nonzero). If `condition` is omitted (in other words, the string is empty) then the optimization is applied to every sequence of insns that matches the patterns.

The defined peephole optimizations are applied after register allocation is complete. Therefore, the peephole definition can check which operands have ended up in which kinds of registers, just by looking at the operands.

The way to refer to the operands in `condition` is to write `operands[i]` for operand number `i` (as matched by `(match_operand i ...)`). Use the variable `insn` to refer to the last of the insns being matched; use `prev_active_insn` to find the preceding insns.

When optimizing computations with intermediate results, you can use `condition` to match only when the intermediate results are not used elsewhere. Use the C expression `dead_or_set_p (insn, op)`, where `insn` is the insn in which you expect the value to be used for the last time (from the value of `insn`, together with use of `prev_nonnote_insn`), and `op` is the intermediate value (from `operands[i]`).

Applying the optimization means replacing the sequence of insns with one new insn. The template controls ultimate output of assembler code for this combined insn. It works exactly like the template of a `define_insn`. Operand numbers in this template are the same ones used in matching the original sequence of insns.

The result of a defined peephole optimizer does not need to match any of the insn patterns in the machine description; it does not even have an opportunity to match them. The peephole optimizer definition itself serves as the insn pattern to control how the insn is output.

Defined peephole optimizers are run as assembler code is being output, so the insns they produce are never combined or rearranged in any way.

Here is an example, taken from the 68000 machine description:

```c
(define_peephole
  [(set (reg:SI 15) (plus:SI (reg:SI 15) (const_int 4)))
   (set (match_operand:DF 0 "register_operand" ";f")
        (match_operand:DF 1 "register_operand" ";ad"))]
  "FP_REG_P (operands[0]) && ! FP_REG_P (operands[1])"
{
  rtx xoperands[2];
  xoperands[1] = gen_rtx_REG (SImode, REGNO (operands[1]) + 1);
  #ifdef MOTOROLA
  output_asm_insn ("move.l %1,(sp)", xoperands);
  output_asm_insn ("move.l %1,-(sp)", operands);
  return "fmove.d (sp)+,%0";
  #else
  output_asm_insn ("movel %1,sp@", xoperands);
  output_asm_insn ("movel %1,sp@-", operands);
  return "fmoved sp@+,%0";
  #endif
}
```
The effect of this optimization is to change

```c
jbsr _foobar
addql #4,sp
movel d1,sp@
movel d0,sp@-
fmoved sp@+,fp0
```
into

```c
jbsr _foobar
movel d1,sp@
movel d0,sp@-
fmoved sp@+,fp0
```

`insn-pattern-1` and so on look *almost* like the second operand of `define_insn`. There is one important difference: the second operand of `define_insn` consists of one or more RTX's enclosed in square brackets. Usually, there is only one: then the same action can be written as an element of a `define_peephole`. But when there are multiple actions in a `define_insn`, they are implicitly enclosed in a `parallel`. Then you must explicitly write the `parallel`, and the square brackets within it, in the `define_peephole`. Thus, if an insn pattern looks like this,

```c
(define_insn "divmodsi4"
[(set (match_operand:SI 0 "general_operand" "=d")
  (div:SI (match_operand:SI 1 "general_operand" "0")
    (match_operand:SI 2 "general_operand" "dmsK")))]
"TARGET_68020"
"divsi%.l %2,%3:%0")
```
then the way to mention this insn in a peephole is as follows:

```c
(define_peephole
[...
(parallel
[(set (match_operand:SI 0 "general_operand" "=d")
  (div:SI (match_operand:SI 1 "general_operand" "0")
    (match_operand:SI 2 "general_operand" "dmsK")))]
...])
```

### 16.18.2 RTL to RTL Peephole Optimizers

The `define_peephole2` definition tells the compiler how to substitute one sequence of instructions for another sequence, what additional scratch registers may be needed and what their lifetimes must be.

```c
(define_peephole2
[insn-pattern-1
 insn-pattern-2
...]
"condition"
[new-insn-pattern-1
 new-insn-pattern-2
...]
```
"preparation-statements")

The definition is almost identical to `define_split` (see Section 16.16 [Insn Splitting], page 364) except that the pattern to match is not a single instruction, but a sequence of instructions.

It is possible to request additional scratch registers for use in the output template. If appropriate registers are not free, the pattern will simply not match.

Scratch registers are requested with a `match_scratch` pattern at the top level of the input pattern. The allocated register (initially) will be dead at the point requested within the original sequence. If the scratch is used at more than a single point, a `match_dup` pattern at the top level of the input pattern marks the last position in the input sequence at which the register must be available.

Here is an example from the IA-32 machine description:

```plaintext
(define_peephole2
  [(match_scratch:SI 2 "r")
   (parallel [(set (match_operand:SI 0 "register_operand" "")
                  (match_operand:SI 3 "arith_or_logical_operator"
                    (match_dup 0)
                    (match_operand:SI 1 "memory_operand" "")))]
              (clobber (reg:CC 17))))]
"! optimize_size && ! TARGET_READ_MODIFY"
[(set (match_dup 2) (match_dup 1))
 (parallel [(set (match_dup 0)
                (match_op_dup 3 [(match_dup 0) (match_dup 2)])
                (clobber (reg:CC 17)))]
          "")
]
```

This pattern tries to split a load from its use in the hopes that we'll be able to schedule around the memory load latency. It allocates a single SImode register of class `GENERAL_REGS` ("r") that needs to be live only at the point just before the arithmetic.

A real example requiring extended scratch lifetimes is harder to come by, so here’s a silly made-up example:

```plaintext
(define_peephole2
  [(match_scratch:SI 4 "r")
   (set (match_operand:SI 0 "")
        (match_operand:SI 1 ""))
   (set (match_operand:SI 2 ""
         (match_dup 1))
        (match_dup 4)
        (set (match_operand:SI 3 "" (match_dup 1))]
 "* determine 1 does not overlap 0 and 2 *
 [(set (match_dup 4) (match_dup 1))
  (set (match_dup 0) (match_dup 4))
  (set (match_dup 2) (match_dup 4))]
 (set (match_dup 3) (match_dup 4))]
"
```

If we had not added the `(match_dup 4)` in the middle of the input sequence, it might have been the case that the register we chose at the beginning of the sequence is killed by the first or second `set`.

### 16.19 Instruction Attributes

In addition to describing the instruction supported by the target machine, the ‘md’ file also defines a group of `attributes` and a set of values for each. Every generated insn is assigned
a value for each attribute. One possible attribute would be the effect that the insn has on the machine’s condition code. This attribute can then be used by NOTICE_UPDATE_CC to track the condition codes.

16.19.1 Defining Attributes and their Values

The `define_attr` expression is used to define each attribute required by the target machine. It looks like:

```
(define_attr name list-of-values default)
```

- `name` is a string specifying the name of the attribute being defined.
- `list-of-values` is either a string that specifies a comma-separated list of values that can be assigned to the attribute, or a null string to indicate that the attribute takes numeric values.
- `default` is an attribute expression that gives the value of this attribute for insns that match patterns whose definition does not include an explicit value for this attribute. See Section 16.19.4 [Attr Example], page 377, for more information on the handling of defaults. See Section 16.19.6 [Constant Attributes], page 379, for information on attributes that do not depend on any particular insn.

For each defined attribute, a number of definitions are written to the ` insn-attr.h` file. For cases where an explicit set of values is specified for an attribute, the following are defined:

- A `#define` is written for the symbol `HAVE_ATTR_name`.
- An enumerated class is defined for `attr_name` with elements of the form `upper-name_upper-value` where the attribute name and value are first converted to upper-case.
- A function `get_attr_name` is defined that is passed an insn and returns the attribute value for that insn.

For example, if the following is present in the `md` file:

```
(define_attr "type" "branch,fp,load,store,arith" ...)
```

the following lines will be written to the file ` insn-attr.h`.

```
#define HAVE_ATTR_type
enum attr_type {TYPE_BRANCH, TYPE_FP, TYPE_LOAD, TYPE_STORE, TYPE_ARITH};
extern enum attr_type get_attr_type ();
```

If the attribute takes numeric values, no `enum` type will be defined and the function to obtain the attribute’s value will return `int`.

There are attributes which are tied to a specific meaning. These attributes are not free to use for other purposes:

- **length** The `length` attribute is used to calculate the length of emitted code chunks. This is especially important when verifying branch distances. See Section 16.19.5 [Insn Lengths], page 378.

- **enabled** The `enabled` attribute can be defined to prevent certain alternatives of an insn definition from being used during code generation. See Section 16.8.6 [Disable Insn Alternatives], page 328.
Another way of defining an attribute is to use:

```
(define_enum_attr "attr" "enum" default)
```

This works in just the same way as `define_attr`, except that the list of values is taken from a separate enumeration called `enum` (see `define_enum`, page 388). This form allows you to use the same list of values for several attributes without having to repeat the list each time. For example:

```
(define_enum "processor" [model_a model_b ...])
(define_enum_attr "arch" "processor"
  (const (symbol_ref "target_arch")))
(define_enum_attr "tune" "processor"
  (const (symbol_ref "target_tune")))
```

defines the same attributes as:

```
(define_attr "arch" "model_a,model_b,..."
  (const (symbol_ref "target_arch")))
(define_attr "tune" "model_a,model_b,..."
  (const (symbol_ref "target_tune")))
```

but without duplicating the processor list. The second example defines two separate C enums (`attr_arch` and `attr_tune`) whereas the first defines a single C enum (`processor`).

### 16.19.2 Attribute Expressions

RTL expressions used to define attributes use the codes described above plus a few specific to attribute definitions, to be discussed below. Attribute value expressions must have one of the following forms:

1. **(const_int i)**
   - The integer `i` specifies the value of a numeric attribute. `i` must be non-negative.
   - The value of a numeric attribute can be specified either with a `const_int`, or as an integer represented as a string in `const_string`, `eq_attr` (see below), `attr`, `symbol_ref`, simple arithmetic expressions, and `set_attr` overrides on specific instructions (see Section 16.19.3 [Tagging Insns], page 376).

2. **(const_string value)**
   - The string `value` specifies a constant attribute value. If `value` is specified as `'"*"' , it means that the default value of the attribute is to be used for the insn containing this expression. `'"*"' obviously cannot be used in the `default` expression of a `define_attr`.
   - If the attribute whose value is being specified is numeric, `value` must be a string containing a non-negative integer (normally `const_int` would be used in this case). Otherwise, it must contain one of the valid values for the attribute.

3. **(if_then_else test true-value false-value)**
   - `test` specifies an attribute test, whose format is defined below. The value of this expression is `true-value` if `test` is true, otherwise it is `false-value`.


\texttt{\textbf{(cond }[\texttt{test1 value1 \ldots} ] \texttt{default)})}

The first operand of this expression is a vector containing an even number of expressions and consisting of pairs of \texttt{test} and \texttt{value} expressions. The value of the \texttt{cond} expression is that of the \texttt{value} corresponding to the first true \texttt{test} expression. If none of the \texttt{test} expressions are true, the value of the \texttt{cond} expression is that of the \texttt{default} expression.

\textit{test} expressions can have one of the following forms:

\texttt{(const\_int }i)\texttt{)}

This test is true if \texttt{i} is nonzero and false otherwise.

\texttt{(not }test)\texttt{)}

\texttt{(ior test1 test2)}

\texttt{(and test1 test2)}

These tests are true if the indicated logical function is true.

\texttt{(match\_operand:\texttt{m} \texttt{n} \texttt{pred} \texttt{constraints})}

This test is true if operand \texttt{n} of the insn whose attribute value is being determined has mode \texttt{m} (this part of the test is ignored if \texttt{m} is \texttt{VOIDmode}) and the function specified by the string \texttt{pred} returns a nonzero value when passed operand \texttt{n} and mode \texttt{m} (this part of the test is ignored if \texttt{pred} is the null string).

The \texttt{constraints} operand is ignored and should be the null string.

\texttt{(le }arith1 \texttt{ arith2)}

\texttt{(leu }arith1 \texttt{ arith2)}

\texttt{(lt }arith1 \texttt{ arith2)}

\texttt{(ltu }arith1 \texttt{ arith2)}

\texttt{(gt }arith1 \texttt{ arith2)}

\texttt{(gtu }arith1 \texttt{ arith2)}

\texttt{(ge }arith1 \texttt{ arith2)}

\texttt{(geu }arith1 \texttt{ arith2)}

\texttt{(ne }arith1 \texttt{ arith2)}

\texttt{(eq }arith1 \texttt{ arith2)}

These tests are true if the indicated comparison of the two arithmetic expressions is true. Arithmetic expressions are formed with \texttt{plus}, \texttt{minus}, \texttt{mult}, \texttt{div}, \texttt{mod}, \texttt{abs}, \texttt{neg}, \texttt{and}, \texttt{ior}, \texttt{xor}, \texttt{not}, \texttt{ashift}, \texttt{lshiftrt}, and \texttt{ashiftrt} expressions.

\texttt{const\_int} and \texttt{symbol\_ref} are always valid terms (see Section 16.19.5 [Insn Lengths], page 378, for additional forms). \texttt{symbol\_ref} is a string denoting a C expression that yields an \texttt{int} when evaluated by the \texttt{get\_attr\_...} routine. It should normally be a global variable.

\texttt{(eq\_attr name value)}

\textit{name} is a string specifying the name of an attribute.

\textit{value} is a string that is either a valid value for attribute \textit{name}, a comma-separated list of values, or \texttt{‘!’} followed by a value or list. If \textit{value} does not begin with a \texttt{‘!’}, this test is true if the value of the \textit{name} attribute of the current insn is in the list specified by \textit{value}. If \textit{value} begins with a \texttt{‘!’}, this test is true if the attribute’s value is \textit{not} in the specified list.
For example,

(eq_attr "type" "load,store")

is equivalent to

(ior (eq_attr "type" "load") (eq_attr "type" "store"))

If name specifies an attribute of 'alternative', it refers to the value of the compiler variable which_alternative (see Section 16.6 [Output Statement], page 296) and the values must be small integers. For example,

(eq_attr "alternative" "2,3")

is equivalent to

(ior (eq (symbol_ref "which_alternative") (const_int 2))
     (eq (symbol_ref "which_alternative") (const_int 3)))

Note that, for most attributes, an eq_attr test is simplified in cases where the value of the attribute being tested is known for all insns matching a particular pattern. This is by far the most common case.

(attr_flag name)

The value of an attr_flag expression is true if the flag specified by name is true for the insn currently being scheduled.

name is a string specifying one of a fixed set of flags to test. Test the flags forward and backward to determine the direction of a conditional branch. Test the flags very_likely, likely, very_unlikely, and unlikely to determine if a conditional branch is expected to be taken.

If the very_likely flag is true, then the likely flag is also true. Likewise for the very_unlikely and unlikely flags.

This example describes a conditional branch delay slot which can be nullified for forward branches that are taken (annul-true) or for backward branches which are not taken (annul-false).

(define_delay (eq_attr "type" "cbranch")
  [(eq_attr "in_branch_delay" "true")
   (and (eq_attr "in_branch_delay" "true")
        (attr_flag "forward"))
   (and (eq_attr "in_branch_delay" "true")
        (attr_flag "backward"))])

The forward and backward flags are false if the current insn being scheduled is not a conditional branch.

The very_likely and likely flags are true if the insn being scheduled is not a conditional branch. The very_unlikely and unlikely flags are false if the insn being scheduled is not a conditional branch.

attr_flag is only used during delay slot scheduling and has no meaning to other passes of the compiler.

(attr name)

The value of another attribute is returned. This is most useful for numeric attributes, as eq_attr and attr_flag produce more efficient code for non-numeric attributes.
16.19.3 Assigning Attribute Values to Insns

The value assigned to an attribute of an insn is primarily determined by which pattern is matched by that insn (or which define_peephole generated it). Every define_insn and define_peephole can have an optional last argument to specify the values of attributes for matching insns. The value of any attribute not specified in a particular insn is set to the default value for that attribute, as specified in its define_attr. Extensive use of default values for attributes permits the specification of the values for only one or two attributes in the definition of most insn patterns, as seen in the example in the next section.

The optional last argument of define_insn and define_peephole is a vector of expressions, each of which defines the value for a single attribute. The most general way of assigning an attribute’s value is to use a set expression whose first operand is an attr expression giving the name of the attribute being set. The second operand of the set is an attribute expression (see Section 16.19.2 [Expressions], page 373) giving the value of the attribute.

When the attribute value depends on the ‘alternative’ attribute (i.e., which is the applicable alternative in the constraint of the insn), the set_attr_alternative expression can be used. It allows the specification of a vector of attribute expressions, one for each alternative.

When the generality of arbitrary attribute expressions is not required, the simpler set_attr expression can be used, which allows specifying a string giving either a single attribute value or a list of attribute values, one for each alternative.

The form of each of the above specifications is shown below. In each case, name is a string specifying the attribute to be set.

(set_attr name value-string)

value-string is either a string giving the desired attribute value, or a string containing a comma-separated list giving the values for succeeding alternatives. The number of elements must match the number of alternatives in the constraint of the insn pattern.

Note that it may be useful to specify ‘*’ for some alternative, in which case the attribute will assume its default value for insns matching that alternative.

(set_attr_alternative name [value1 value2 ...])

Depending on the alternative of the insn, the value will be one of the specified values. This is a shorthand for using a cond with tests on the ‘alternative’ attribute.

(set (attr name) value)

The first operand of this set must be the special RTL expression attr, whose sole operand is a string giving the name of the attribute being set. value is the value of the attribute.

The following shows three different ways of representing the same attribute value specification:

(set_attr "type" "load,store,arith")

(set_attr_alternative "type"
    [(const_string "load") (const_string "store")]

define_asm_attributes

The define_asm_attributes expression provides a mechanism to specify the attributes assigned to insns produced from an asm statement. It has the form:

(define_asm_attributes [attr-sets])

where attr-sets is specified the same as for both the define_insn and the define_peephole expressions.

These values will typically be the “worst case” attribute values. For example, they might indicate that the condition code will be clobbered.

A specification for a length attribute is handled specially. The way to compute the length of an asm insn is to multiply the length specified in the expression define_asm_attributes by the number of machine instructions specified in the asm statement, determined by counting the number of semicolons and newlines in the string. Therefore, the value of the length attribute specified in a define_asm_attributes should be the maximum possible length of a single machine instruction.

16.19.4 Example of Attribute Specifications

The judicious use of defaulting is important in the efficient use of insn attributes. Typically, insns are divided into types and an attribute, customarily called type, is used to represent this value. This attribute is normally used only to define the default value for other attributes. An example will clarify this usage.

Assume we have a RISC machine with a condition code and in which only full-word operations are performed in registers. Let us assume that we can divide all insns into loads, stores, (integer) arithmetic operations, floating point operations, and branches.

Here we will concern ourselves with determining the effect of an insn on the condition code and will limit ourselves to the following possible effects: The condition code can be set unpredictably (clobbered), not be changed, be set to agree with the results of the operation, or only changed if the item previously set into the condition code has been modified.

Here is part of a sample ‘md’ file for such a machine:

(define_attr "type" "load,store,arith,fp,branch" (const_string "arith"))

(define_attr "cc" "clobber,unchanged,set,change0"
  (cond [(eq_attr "type" "load")
          (const_string "change0")
          (eq_attr "type" "store,branch")
          (const_string "unchanged")
          (eq_attr "type" "arith")
          (if_then_else (match_operand:SI 0 "" ")
          (const_string "set")
          (const_string "clobber"))]
      (const_string "clobber"))

(define_insn ""
  [(set (match_operand:SI 0 "general_operand" "=r,r,m")
       (match_operand:SI 1 "general_operand" "r,m,r"))]
"@ move %0,%1
load %0,%1
store %0,%1"

[(set_attr "type" "arith,load,store")]

Note that we assume in the above example that arithmetic operations performed on quantities smaller than a machine word clobber the condition code since they will set the condition code to a value corresponding to the full-word result.

16.19.5 Computing the Length of anInsn

For many machines, multiple types of branch instructions are provided, each for different length branch displacements. In most cases, the assembler will choose the correct instruction to use. However, when the assembler cannot do so, GCC can do it when a special attribute, the length attribute, is defined. This attribute must be defined to have numeric values by specifying a null string in its define_attr.

In the case of the length attribute, two additional forms of arithmetic terms are allowed in test expressions:

(match_dup n)
This refers to the address of operand n of the current insn, which must be a label_ref.

(pc)
This refers to the address of the current insn. It might have been more consistent with other usage to make this the address of the next insn but this would be confusing because the length of the current insn is to be computed.

For normal insns, the length will be determined by value of the length attribute. In the case of addr_vec and addr_diff_vec insn patterns, the length is computed as the number of vectors multiplied by the size of each vector.

Lengths are measured in addressable storage units (bytes).

The following macros can be used to refine the length computation:

ADJUST_INSN_LENGTH (insn, length)
If defined, modifies the length assigned to instruction insn as a function of the context in which it is used. length is an lvalue that contains the initially computed length of the insn and should be updated with the correct length of the insn.

This macro will normally not be required. A case in which it is required is the ROMP. On this machine, the size of an addr_vec insn must be increased by two to compensate for the fact that alignment may be required.

The routine that returns get_attr_length (the value of the length attribute) can be used by the output routine to determine the form of the branch instruction to be written, as the example below illustrates.

As an example of the specification of variable-length branches, consider the IBM 360. If we adopt the convention that a register will be set to the starting address of a function, we can jump to labels within 4k of the start using a four-byte instruction. Otherwise, we need a six-byte sequence to load the address from memory and then branch to it.

On such a machine, a pattern for a branch instruction might be specified as follows:
(define_insn "jump"

  [(set (pc)
    (label_ref (match_operand 0 "" "")))]

  ""

  {
    return (get_attr_length (insn) == 4
      ? "b %l0" : "l r15,=a(%l0); br r15");
  }

  [(set (attr "length")
    (if_then_else (lt (match_dup 0) (const_int 4096))
      (const_int 4)
      (const_int 6)))]

16.19.6 Constant Attributes

A special form of define_attr, where the expression for the default value is a const
expression, indicates an attribute that is constant for a given run of the compiler. Constant
attributes may be used to specify which variety of processor is used. For example,

  (define_attr "cpu" "m88100,m88110,m88000"
    (const
      (cond [(symbol_ref "TARGET_88100") (const_string "m88100")
            (symbol_ref "TARGET_88110") (const_string "m88110")
            (const_string "m88000"))]))

  (define_attr "memory" "fast,slow"
    (const
      (if_then Else (symbol_ref "TARGET_FAST_MEM")
        (const_string "fast")
        (const_string "slow"))))

The routine generated for constant attributes has no parameters as it does not depend
on any particular insn. RTL expressions used to define the value of a constant attribute
may use the symbol_ref form, but may not use either the match_operand form or eq_attr
forms involving insn attributes.

16.19.7 Delay Slot Scheduling

The insn attribute mechanism can be used to specify the requirements for delay slots, if any,
on a target machine. An instruction is said to require a delay slot if some instructions that
are physically after the instruction are executed as if they were located before it. Classic
eamples are branch and call instructions, which often execute the following instruction
before the branch or call is performed.

On some machines, conditional branch instructions can optionally annul instructions in
the delay slot. This means that the instruction will not be executed for certain branch
outcomes. Both instructions that annul if the branch is true and instructions that annul if
the branch is false are supported.

Delay slot scheduling differs from instruction scheduling in that determining whether an
instruction needs a delay slot is dependent only on the type of instruction being generated,
not on data flow between the instructions. See the next section for a discussion of data-
dependent instruction scheduling.

The requirement of an insn needing one or more delay slots is indicated via the define_-
delay expression. It has the following form:

  (define_delay test
test is an attribute test that indicates whether this define_delay applies to a particular insn. If so, the number of required delay slots is determined by the length of the vector specified as the second argument. An insn placed in delay slot n must satisfy attribute test delay-n. annul-true-n is an attribute test that specifies which insns may be annulled if the branch is true. Similarly, annul-false-n specifies which insns in the delay slot may be annulled if the branch is false. If annulling is not supported for that delay slot, (nil) should be coded.

For example, in the common case where branch and call insns require a single delay slot, which may contain any insn other than a branch or call, the following would be placed in the 'md' file:

```lisp
(define_delay (eq_attr "type" "branch,call")
  [(eq_attr "type" "!branch,call") (nil) (nil)])
```

Multiple define_delay expressions may be specified. In this case, each such expression specifies different delay slot requirements and there must be no insn for which tests in two define_delay expressions are both true.

For example, if we have a machine that requires one delay slot for branches but two for calls, no delay slot can contain a branch or call insn, and any valid insn in the delay slot for the branch can be annulled if the branch is true, we might represent this as follows:

```lisp
(define_delay (eq_attr "type" "branch")
  [(eq_attr "type" "!branch,call")
   (eq_attr "type" "!branch,call")
   (nil)])

(define_delay (eq_attr "type" "call")
  [(eq_attr "type" "!branch,call") (nil) (nil)
   (eq_attr "type" "!branch,call") (nil) (nil)])
```

16.19.8 Specifying processor pipeline description

To achieve better performance, most modern processors (super-pipelined, superscalar RISC, and VLIW processors) have many functional units on which several instructions can be executed simultaneously. An instruction starts execution if its issue conditions are satisfied. If not, the instruction is stalled until its conditions are satisfied. Such interlock (pipeline) delay causes interruption of the fetching of successor instructions (or demands nop instructions, e.g. for some MIPS processors).

There are two major kinds of interlock delays in modern processors. The first one is a data dependence delay determining instruction latency time. The instruction execution is not started until all source data have been evaluated by prior instructions (there are more complex cases when the instruction execution starts even when the data are not available but will be ready in given time after the instruction execution start). Taking the data dependence delays into account is simple. The data dependence (true, output, and anti-dependence) delay between two instructions is given by a constant. In most cases this approach is adequate. The second kind of interlock delays is a reservation delay. The reservation delay means that two instructions under execution will be in need of shared processors resources, i.e. buses, internal registers, and/or functional units, which are reserved
for some time. Taking this kind of delay into account is complex especially for modern RISC processors.

The task of exploiting more processor parallelism is solved by an instruction scheduler. For a better solution to this problem, the instruction scheduler has to have an adequate description of the processor parallelism (or pipeline description). GCC machine descriptions describe processor parallelism and functional unit reservations for groups of instructions with the aid of regular expressions.

The GCC instruction scheduler uses a pipeline hazard recognizer to figure out the possibility of the instruction issue by the processor on a given simulated processor cycle. The pipeline hazard recognizer is automatically generated from the processor pipeline description. The pipeline hazard recognizer generated from the machine description is based on a deterministic finite state automaton (DFA): the instruction issue is possible if there is a transition from one automaton state to another one. This algorithm is very fast, and furthermore, its speed is not dependent on processor complexity.

The rest of this section describes the directives that constitute an automaton-based processor pipeline description. The order of these constructions within the machine description file is not important.

The following optional construction describes names of automata generated and used for the pipeline hazards recognition. Sometimes the generated finite state automaton used by the pipeline hazard recognizer is large. If we use more than one automaton and bind functional units to the automata, the total size of the automata is usually less than the size of the single automaton. If there is no such construction, only one finite state automaton is generated.

\[
\text{(define_automaton automata-names)}
\]

*automata-names* is a string giving names of the automata. The names are separated by commas. All the automata should have unique names. The automaton name is used in the constructions *define_cpu_unit* and *define_query_cpu_unit*.

Each processor functional unit used in the description of instruction reservations should be described by the following construction.

\[
\text{(define_cpu_unit unit-names [automaton-name])}
\]

*unit-names* is a string giving the names of the functional units separated by commas. Don’t use name ‘nothing’, it is reserved for other goals.

*automaton-name* is a string giving the name of the automaton with which the unit is bound. The automaton should be described in construction *define_automaton*. You should give *automaton-name*, if there is a defined automaton.

The assignment of units to automata are constrained by the uses of the units in insn reservations. The most important constraint is: if a unit reservation is present on a particular cycle of an alternative for an insn reservation, then some unit from the same automaton must be present on the same cycle for the other alternatives of the insn reservation. The rest of the constraints are mentioned in the description of the subsequent constructions.

---

2 However, the size of the automaton depends on processor complexity. To limit this effect, machine descriptions can split orthogonal parts of the machine description among several automata: but then, since each of these must be stepped independently, this does cause a small decrease in the algorithm’s performance.
The following construction describes CPU functional units analogously to\texttt{define_cpu_unit}. The reservation of such units can be queried for an automaton state. The instruction scheduler never queries reservation of functional units for given automaton state. So as a rule, you don’t need this construction. This construction could be used for future code generation goals (e.g. to generate VLIW insn templates).

\begin{verbatim}
(define_query_cpu_unit unit-names [automaton-name])
\end{verbatim}

\texttt{unit-names} is a string giving names of the functional units separated by commas.

\texttt{automaton-name} is a string giving the name of the automaton with which the unit is bound.

The following construction is the major one to describe pipeline characteristics of an instruction.

\begin{verbatim}
(define_insn_reservation insn-name default_latency condition regexp)
\end{verbatim}

\texttt{default_latency} is a number giving latency time of the instruction. There is an important difference between the old description and the automaton based pipeline description. The latency time is used for all dependencies when we use the old description. In the automaton based pipeline description, the given latency time is only used for true dependencies. The cost of anti-dependencies is always zero and the cost of output dependencies is the difference between latency times of the producing and consuming insns (if the difference is negative, the cost is considered to be zero). You can always change the default costs for any description by using the target hook \texttt{TARGET_SCHED_ADJUST_COST} (see Section 17.18 [Scheduling], page 485).

\texttt{insn-name} is a string giving the internal name of the insn. The internal names are used in constructions \texttt{define_bypass} and in the automaton description file generated for debugging. The internal name has nothing in common with the names in \texttt{define_insn}. It is a good practice to use insn classes described in the processor manual.

\texttt{condition} defines what RTL insns are described by this construction. You should remember that you will be in trouble if \texttt{condition} for two or more different \texttt{define_insn_reservation} constructions is TRUE for an insn. In this case what reservation will be used for the insn is not defined. Such cases are not checked during generation of the pipeline hazards recognizer because in general recognizing that two conditions may have the same value is quite difficult (especially if the conditions contain \texttt{symbol_ref}). It is also not checked during the pipeline hazard recognizer work because it would slow down the recognizer considerably.

\texttt{regexp} is a string describing the reservation of the cpu’s functional units by the instruction. The reservations are described by a regular expression according to the following syntax:

\begin{verbatim}
regexp = regexp "," oneof
  | oneof

oneof = oneof "|" allof
  | allof

allof = allof "*" repeat
  | repeat

repeat = element "*" number
\end{verbatim}
Chapter 16: Machine Descriptions

| element
| element = cpu_function_unit_name
| reservation_name
| result_name
| "nothing"
| "(" regexp ")"

- `|` is used for describing the start of the next cycle in the reservation.
- `|` is used for describing a reservation described by the first regular expression or a reservation described by the second regular expression or etc.
- `+` is used for describing a reservation described by the first regular expression and a reservation described by the second regular expression and etc.
- `*` is used for convenience and simply means a sequence in which the regular expression are repeated number times with cycle advancing (see `|`).
- `cpu_function_unit_name` denotes reservation of the named functional unit.
- `reservation_name` — see description of construction `define_reservation`.
- `nothing` denotes no unit reservations.

Sometimes unit reservations for different insns contain common parts. In such case, you can simplify the pipeline description by describing the common part by the following construction

\[
\text{(define_reservation reservation-name regexp)}
\]

`reservation-name` is a string giving name of regexp. Functional unit names and reservation names are in the same name space. So the reservation names should be different from the functional unit names and can not be the reserved name `nothing`.

The following construction is used to describe exceptions in the latency time for given instruction pair. This is so called bypasses.

\[
\text{(define_bypass number out_insn_names in_insn_names [guard])}
\]

`number` defines when the result generated by the instructions given in string `out_insn_names` will be ready for the instructions given in string `in_insn_names`. The instructions in the string are separated by commas.

`guard` is an optional string giving the name of a C function which defines an additional guard for the bypass. The function will get the two insns as parameters. If the function returns zero the bypass will be ignored for this case. The additional guard is necessary to recognize complicated bypasses, e.g. when the consumer is only an address of insn `store` (not a stored value).

If there are more one bypass with the same output and input insns, the chosen bypass is the first bypass with a guard in description whose guard function returns nonzero. If there is no such bypass, then bypass without the guard function is chosen.

The following five constructions are usually used to describe VLIW processors, or more precisely, to describe a placement of small instructions into VLIW instruction slots. They can be used for RISC processors, too.

\[
\text{(exclusion_set unit-names unit-names)}
\]
\[
\text{(presence_set unit-names patterns)}
\]
\[
\text{(final_presence_set unit-names patterns)}
\]
unit-names is a string giving names of functional units separated by commas.

patterns is a string giving patterns of functional units separated by comma. Currently pattern is one unit or units separated by white-spaces.

The first construction (‘exclusion_set’) means that each functional unit in the first string can not be reserved simultaneously with a unit whose name is in the second string and vice versa. For example, the construction is useful for describing processors (e.g. some SPARC processors) with a fully pipelined floating point functional unit which can execute simultaneously only single floating point insns or only double floating point insns.

The second construction (‘presence_set’) means that each functional unit in the first string can not be reserved unless at least one of pattern of units whose names are in the second string is reserved. This is an asymmetric relation. For example, it is useful for description that VLIW ‘slot1’ is reserved after ‘slot0’ reservation. We could describe it by the following construction

(absence_set unit-names patterns)
(final_absence_set unit-names patterns)

The forth construction (‘absence_set’) means that each functional unit in the first string can be reserved only if each pattern of units whose names are in the second string is not reserved. This is an asymmetric relation (actually ‘exclusion_set’ is analogous to this one but it is symmetric). For example it might be useful in a VLIW description to say that ‘slot0’ cannot be reserved after either ‘slot1’ or ‘slot2’ have been reserved. This can be described as:

(absence_set unit-names patterns)
(absence_set "slot0" "slot1, slot2")

All functional units mentioned in a set should belong to the same automaton.

The last construction (‘final_absence_set’) is analogous to ‘absence_set’ but checking is done on the result (state) reservation. See comments for ‘final_presence_set’.
You can control the generator of the pipeline hazard recognizer with the following construction.

```
(automata-option options)
```

`options` is a string giving options which affect the generated code. Currently there are the following options:

- **no-minimization** makes no minimization of the automaton. This is only worth to do when we are debugging the description and need to look more accurately at reservations of states.
- **time** means printing time statistics about the generation of automata.
- **stats** means printing statistics about the generated automata such as the number of DFA states, NDFA states and arcs.
- **v** means a generation of the file describing the result automata. The file has suffix `.dfa` and can be used for the description verification and debugging.
- **w** means a generation of warning instead of error for non-critical errors.
- **ndfa** makes nondeterministic finite state automata. This affects the treatment of operator `'|'` in the regular expressions. The usual treatment of the operator is to try the first alternative and, if the reservation is not possible, the second alternative. The nondeterministic treatment means trying all alternatives, some of them may be rejected by reservations in the subsequent insns.
- **progress** means output of a progress bar showing how many states were generated so far for automaton being processed. This is useful during debugging a DFA description. If you see too many generated states, you could interrupt the generator of the pipeline hazard recognizer and try to figure out a reason for generation of the huge automaton.

As an example, consider a superscalar RISC machine which can issue three insns (two integer insns and one floating point insn) on the cycle but can finish only two insns. To describe this, we define the following functional units.

```
(define-cpu-unit "i0_pipeline, i1_pipeline, f_pipeline")
(define-cpu-unit "port0, port1")
```

All simple integer insns can be executed in any integer pipeline and their result is ready in two cycles. The simple integer insns are issued into the first pipeline unless it is reserved, otherwise they are issued into the second pipeline. Integer division and multiplication insns can be executed only in the second integer pipeline and their results are ready correspondingly in 8 and 4 cycles. The integer division is not pipelined, i.e. the subsequent integer division insn can not be issued until the current division insn finished. Floating point insns are fully pipelined and their results are ready in 3 cycles. Where the result of a floating point insn is used by an integer insn, an additional delay of one cycle is incurred. To describe all of this we could specify

```
(define-cpu-unit "div")

(define-instruction-reservation "simple" 2 (eq_attr "type" "int")
    "(i0_pipeline | i1_pipeline), (port0 | port1)"
)

(define-instruction-reservation "mult" 4 (eq_attr "type" "mult")
    "i1_pipeline, nothing*2, (port0 | port1)"
)

(define-instruction-reservation "div" 8 (eq_attr "type" "div")
)
"i1_pipeline, div7, div + (port0 | port1)"

(define_insn_reservation "float" 3 (eq_attr "type" "float"
"f_pipeline, nothing, (port0 | port1))

(define_bypass 4 "float" "simple,mult,div")

To simplify the description we could describe the following reservation

(define_reservation "finish" "port0|port1")

and use it in all define_insn_reservation as in the following construction

(define_insn_reservation "simple" 2 (eq_attr "type" "int"
"(i0_pipeline | i1_pipeline), finish")

16.20 Conditional Execution

A number of architectures provide for some form of conditional execution, or predication.
The hallmark of this feature is the ability to nullify most of the instructions in the instruction
set. When the instruction set is large and not entirely symmetric, it can be quite tedious
to describe these forms directly in the `.md' file. An alternative is the define_cond_exec
template.

(define_cond_exec
[predicate-pattern]
"condition"
"output-template")

**predicate-pattern** is the condition that must be true for the insn to be executed at runtime
and should match a relational operator. One can use **match_operator** to match several
relational operators at once. Any **match_operand** operands must have no more than one
alternative.

**condition** is a C expression that must be true for the generated pattern to match.

**output-template** is a string similar to the define_insn output template (see Section 16.5
[Output Template], page 295), except that the ‘*’ and ‘@’ special cases do not apply. This
is only useful if the assembly text for the predicate is a simple prefix to the main insn. In
order to handle the general case, there is a global variable **current_insn_predicate** that
will contain the entire predicate if the current insn is predicated, and will otherwise be NULL.

When define_cond_exec is used, an implicit reference to the **predicable** instruction
attribute is made. See Section 16.19 [Insn Attributes], page 371. This attribute must be
boolean (i.e. have exactly two elements in its **list-of-values**). Further, it must not be used
with complex expressions. That is, the default and all uses in the insns must be a simple
constant, not dependent on the alternative or anything else.

For each define_insn for which the **predicable** attribute is true, a new define_insn
pattern will be generated that matches a predicated version of the instruction. For example,

(define_insn "addsi"
[(set (match_operand:SI 0 "register_operand" "r")
 (plus:SI (match_operand:SI 1 "register_operand" "r")
  (match_operand:SI 2 "register_operand" "r"))))]
"test1"
"add %2,%1,%0")

(define_cond_exec
[(ne (match_operand:CC 0 "register_operand" "c")
generates a new pattern

\[
\text{(define_insn ""
  [(cond_exec
    (ne (match_operand:CC 3 "register_operand" "c") (const_int 0))
    (set (match_operand:SI 0 "register_operand" "r")
      (plus:SI (match_operand:SI 1 "register_operand" "r")
        (match_operand:SI 2 "register_operand" "r"))))
  "((test2) && (test1))"
  "(%3) add %2,%1,%0")}
\]

16.21 Constant Definitions

Using literal constants inside instruction patterns reduces legibility and can be a mainte-
nance problem.

To overcome this problem, you may use the \texttt{define_constants} expression. It contains
a vector of name-value pairs. From that point on, wherever any of the names appears in
the MD file, it is as if the corresponding value had been written instead. You may use
\texttt{define_constants} multiple times; each appearance adds more constants to the table. It
is an error to redefine a constant with a different value.

To come back to the a29k load multiple example, instead of

\[
\text{(define_insn ""
  [(match_parallel 0 "load_multiple_operation"
    [(set (match_operand:SI 1 "gpc_reg_operand" ")=r")(match_operand:SI 2 "memory_operand" "m")
    (use (reg:SI 179))
    (clobber (reg:SI 179)))]
  "" "loadm 0,0,%1,%2")}
\]

You could write:

\[
\text{(define_constants [}
  (R_BP 177)
  (R_FC 178)
  (R_CR 179)
  (R_Q  180)
])
\]

\[
\text{(define_insn ""
  [(match_parallel 0 "load_multiple_operation"
    [(set (match_operand:SI 1 "gpc_reg_operand" ")=r")
    (match_operand:SI 2 "memory_operand" "m")
    (use (reg:SI R_CR))
    (clobber (reg:SI R_CR)))]
  "" "loadm 0,0,%1,%2")}
\]

The constants that are defined with a \texttt{define_constant} are also output in the insn-codes.h
header file as \#defines.

You can also use the machine description file to define enumerations. Like the constants
defined by \texttt{define_constant}, these enumerations are visible to both the machine description
file and the main C code.
The syntax is as follows:

```c
(define_c_enum "name" [value0 value1 ...
valuenn ]]
)
```

This definition causes the equivalent of the following C code to appear in ‘insn-constants.h’:

```c
enum name {
    value0 = 0,
    value1 = 1,
    ...
    valuenn = n
};
#define NUM_cname_VALUES (n + 1)
```

where `cname` is the capitalized form of `name`. It also makes each `valuei` available in the machine description file, just as if it had been declared with:

```c
(define_constants [(valuei i)])
```

Each `valuei` is usually an upper-case identifier and usually begins with `cname`.

You can split the enumeration definition into as many statements as you like. The above example is directly equivalent to:

```c
(define_c_enum "name" [value0])
(define_c_enum "name" [value1])
...
(define_c_enum "name" [valuenn])
```

Splitting the enumeration helps to improve the modularity of each individual .md file. For example, if a port defines its synchronization instructions in a separate ‘sync.md’ file, it is convenient to define all synchronization-specific enumeration values in ‘sync.md’ rather than in the main .md file.

Some enumeration names have special significance to GCC:

- **unspecv** If an enumeration called `unspecv` is defined, GCC will use it when printing out `unspec_volatile` expressions. For example:
  ```c
  (define_c_enum "unspecv" [UNSPECV_BLOCKAGE])
  ```

  causes GCC to print ‘`(unspec_volatile ... 0)`’ as:

  ```c
  (unspec_volatile ... UNSPECV_BLOCKAGE)
  ```

- **unspec** If an enumeration called `unspec` is defined, GCC will use it when printing out `unspec` expressions. GCC will also use it when printing out `unspec_volatile` expressions unless an `unspecv` enumeration is also defined. You can therefore decide whether to keep separate enumerations for volatile and non-volatile expressions or whether to use the same enumeration for both.

Another way of defining an enumeration is to use `define_enum`:

```c
(define_enum "name" [value0 value1 ...
```
This directive implies:

```
(define_c_enum "name" [
  cname_cvalue0
  cname_cvalue1
  ...
  cname_cvaluen
])
```

where `cvaluei` is the capitalized form of `valuei`. However, unlike `define_c_enum`, the enumerations defined by `define_enum` can be used in attribute specifications (see [define_enum_attr], page 373).

### 16.22 Iterators

Ports often need to define similar patterns for more than one machine mode or for more than one rtx code. GCC provides some simple iterator facilities to make this process easier.

#### 16.22.1 Mode Iterators

Ports often need to define similar patterns for two or more different modes. For example:

- If a processor has hardware support for both single and double floating-point arithmetic, the `SFmode` patterns tend to be very similar to the `DFmode` ones.
- If a port uses `SImode` pointers in one configuration and `DImode` pointers in another, it will usually have very similar `SImode` and `DImode` patterns for manipulating pointers.

Mode iterators allow several patterns to be instantiated from one `.md` file template. They can be used with any type of rtx-based construct, such as a `define_insn`, `define_split`, or `define_peephole2`.

#### 16.22.1.1 Defining Mode Iterators

The syntax for defining a mode iterator is:

```
(define_mode_iterator name [(mode1 "cond1") ... (moden "condn")])
```

This allows subsequent `.md` file constructs to use the mode suffix `:name`. Every construct that does so will be expanded `n` times, once with every use of `:name` replaced by `:mode1`, once with every use replaced by `:mode2`, and so on. In the expansion for a particular `modei`, every C condition will also require that `condi` be true.

For example:

```
(define_mode_iterator P [(SI "Pmode == SImode") (DI "Pmode == DImode")])
```

defines a new mode suffix :P. Every construct that uses :P will be expanded twice, once with every :P replaced by :SI and once with every :P replaced by :DI. The :SI version will only apply if `Pmode == SImode` and the :DI version will only apply if `Pmode == DImode`.

As with other `.md` conditions, an empty string is treated as “always true”. (mode "") can also be abbreviated to `mode`. For example:

```
(define_mode_iterator GPR [SI (DI "TARGET_64BIT")])
```

means that the :DI expansion only applies if `TARGET_64BIT` but that the :SI expansion has no such constraint.
Iterators are applied in the order they are defined. This can be significant if two iterators are used in a construct that requires substitutions. See Section 16.22.1.2 [Substitutions], page 390.

16.22.1.2 Substitution in Mode Iterators

If an `.md` file construct uses mode iterators, each version of the construct will often need slightly different strings or modes. For example:

- When a `define_expand` defines several adm3 patterns (see Section 16.9 [Standard Names], page 332), each expander will need to use the appropriate mode name for m.
- When a `define_insn` defines several instruction patterns, each instruction will often use a different assembler mnemonic.
- When a `define_insn` requires operands with different modes, using an iterator for one of the operand modes usually requires a specific mode for the other operand(s).

GCC supports such variations through a system of “mode attributes”. There are two standard attributes: `mode`, which is the name of the mode in lower case, and `MODE`, which is the same thing in upper case. You can define other attributes using:

```c
(define_mode_attr name [((mode1 "value1") ... ( moden "valuen"))])
```

where `name` is the name of the attribute and `valuei` is the value associated with `modei`.

When GCC replaces some `:iterator` with `:mode`, it will scan each string and mode in the pattern for sequences of the form `<iterator:attr>`, where `attr` is the name of a mode attribute. If the attribute is defined for `mode`, the whole `<...>` sequence will be replaced by the appropriate value.

For example, suppose an `.md` file has:

```c
(define_mode_iterator P [((SI "Pmode == SImode") (DI "Pmode == DImode")])
(define_mode_attr load [((SI "lw") (DI "ld"))])
```

If one of the patterns that uses :P contains the string `<P:load>	%0,%1`, the SI version of that pattern will use "lw\t%0,%1" and the DI version will use "ld\t%0,%1".

Here is an example of using an attribute for a mode:

```c
(define_mode_iterator LONG [SI DI])
(define_mode_attr SHORT [((SI "HI") (DI "SI"))])
(define_insn ...
 (sign_extend:LONG (match_operand:<LONG:SHORT> ...)) ...
)
```

The `iterator`: prefix may be omitted, in which case the substitution will be attempted for every iterator expansion.

16.22.1.3 Mode Iterator Examples

Here is an example from the MIPS port. It defines the following modes and attributes (among others):

```c
(define_mode_iterator GPR [SI (DI "TARGET_64BIT")])
(define_mode_attr d [((SI ")") (DI "d")])
```

and uses the following template to define both subsi3 and subdi3:

```c
(define_insn "sub<mode>3"
 [(set (match_operand:GPR 0 "register_operand" "]=d")
   (minus:GPR (match_operand:GPR 1 "register_operand" "d")
   (match_operand:GPR 2 "register_operand" "d"))])
```

```c
""
```
"<d>subu\%0,\%1,\%2"
     [(set_attr "type" "arith")
      (set_attr "mode" "<MODE>")])

This is exactly equivalent to:

(define_insn "subsi3"
     [(set (match_operand:SI 0 "register_operand" =d")
          (minus:SI (match_operand:SI 1 "register_operand" d")
                    (match_operand:SI 2 "register_operand" d")))]
     ""
     [(set_attr "type" "arith")
      (set_attr "mode" "SI")])

(define_insn "subdi3"
     [(set (match_operand:DI 0 "register_operand" =d")
          (minus:DI (match_operand:DI 1 "register_operand" d")
                    (match_operand:DI 2 "register_operand" d")))]
     ""
     [(set_attr "type" "arith")
      (set_attr "mode" "DI")])

16.22.2 Code Iterators

Code iterators operate in a similar way to mode iterators. See Section 16.22.1 [Mode Iterators], page 389.

The construct:

(define_code_iterator name [(code1 "cond1") ... (coden "condn")])

defines a pseudo rtx code name that can be instantiated as codei if condition condi is true. Each codei must have the same rtx format. See Section 10.2 [RTL Classes], page 126.

As with mode iterators, each pattern that uses name will be expanded n times, once with all uses of name replaced by code1, once with all uses replaced by code2, and so on. See Section 16.22.1.1 [Defining Mode Iterators], page 389.

It is possible to define attributes for codes as well as for modes. There are two standard code attributes: code, the name of the code in lower case, and CODE, the name of the code in upper case. Other attributes are defined using:

(define_code_attr name [(code1 "value1") ... (coden "valuen")])

Here’s an example of code iterators in action, taken from the MIPS port:

(define_code_iterator any_cond [unordered ordered unlt unge uneq ltgt unle ungt eq ne ge lt le gtu geu ltu leu])

(define_expand "b<code>"
     [(set (pc)
           (if_then_else (any_cond:CC (cc0)
                                        (const_int 0))
                         (label_ref (match_operand 0 ""))
                         (pc)))]
     ""
     {
      gen_conditional_branch (operands, <CODE>);
      DONE;
     }
)

This is equivalent to:
(define_expand "bunordered"
[(set (pc)
      (if_then_else (unordered:CC (cc0)
                       (const_int 0))
       (label_ref (match_operand 0 ""))
       (pc)))]]
"
{
  gen_conditional_branch (operands, UNORDERED);
  DONE;
}
)

(define_expand "bordered"
[(set (pc)
      (if_then_else (ordered:CC (cc0)
                       (const_int 0))
       (label_ref (match_operand 0 ""))
       (pc)))]]
"
{
  gen_conditional_branch (operands, ORDERED);
  DONE;
}
)
Chapter 17: Target Description Macros and Functions

In addition to the file `machine.md`, a machine description includes a C header file conventionally given the name `machine.h` and a C source file named `machine.c`. The header file defines numerous macros that convey the information about the target machine that does not fit into the scheme of the `.md` file. The file `tm.h` should be a link to `machine.h`. The header file `config.h` includes `tm.h` and most compiler source files include `config.h`. The source file defines a variable `targetm`, which is a structure containing pointers to functions and data relating to the target machine. `machine.c` should also contain their definitions, if they are not defined elsewhere in GCC, and other functions called through the macros defined in the `.h` file.

17.1 The Global `targetm` Variable

```c
#include "target.h"
#include "target-def.h"

/* Initialize the GCC target structure. */
#undef TARGET_COMP_TYPE_ATTRIBUTES
#define TARGET_COMP_TYPE_ATTRIBUTES

struct gcc_target targetm = TARGET_INITIALIZER;
```

Where a macro should be defined in the `.c` file in this manner to form part of the `targetm` structure, it is documented below as a “Target Hook” with a prototype. Many macros will change in future from being defined in the `.h` file to being part of the `targetm` structure.

17.2 Controlling the Compilation Driver, `gcc`

You can control the compilation driver.

```c
DRIVER_SELF_SPECS
```

A list of specs for the driver itself. It should be a suitable initializer for an array of strings, with no surrounding braces.

The driver applies these specs to its own command line between loading default `specs` files (but not command-line specified ones) and choosing the multilib directory or running any subcommands. It applies them in the order given, so each spec can depend on the options added by earlier ones. It is also possible to remove options using `%<option>` in the usual way.
This macro can be useful when a port has several interdependent target options. It provides a way of standardizing the command line so that the other specs are easier to write.

Do not define this macro if it does not need to do anything.

**OPTION_DEFAULT_SPECS**  
[Macro]  
A list of specs used to support configure-time default options (i.e. ‘--with’ options) in the driver. It should be a suitable initializer for an array of structures, each containing two strings, without the outermost pair of surrounding braces.

The first item in the pair is the name of the default. This must match the code in ‘config.gcc’ for the target. The second item is a spec to apply if a default with this name was specified. The string ‘%(VALUE)’ in the spec will be replaced by the value of the default everywhere it occurs.

The driver will apply these specs to its own command line between loading default ‘specs’ files and processing DRIVER_SELF_SPECS, using the same mechanism as DRIVER_SELF_SPECS.

Do not define this macro if it does not need to do anything.

**CPP_SPEC**  
[Macro]  
A C string constant that tells the GCC driver program options to pass to CPP. It can also specify how to translate options you give to GCC into options for GCC to pass to the CPP.

Do not define this macro if it does not need to do anything.

**CPLUSPLUS_CPP_SPEC**  
[Macro]  
This macro is just like CPP_SPEC, but is used for C++, rather than C. If you do not define this macro, then the value of CPP_SPEC (if any) will be used instead.

**CC1_SPEC**  
[Macro]  
A C string constant that tells the GCC driver program options to pass to cc1, cc1plus, f771, and the other language front ends. It can also specify how to translate options you give to GCC into options for GCC to pass to front ends.

Do not define this macro if it does not need to do anything.

**CC1PLUS_SPEC**  
[Macro]  
A C string constant that tells the GCC driver program options to pass to cc1plus. It can also specify how to translate options you give to GCC into options for GCC to pass to the cc1plus.

Do not define this macro if it does not need to do anything. Note that everything defined in CC1_SPEC is already passed to cc1plus so there is no need to duplicate the contents of CC1_SPEC in CC1PLUS_SPEC.

**ASM_SPEC**  
[Macro]  
A C string constant that tells the GCC driver program options to pass to the assembler. It can also specify how to translate options you give to GCC into options for GCC to pass to the assembler. See the file ‘sun3.h’ for an example of this.

Do not define this macro if it does not need to do anything.
**ASM_FINAL_SPEC**

A C string constant that tells the GCC driver program how to run any programs which cleanup after the normal assembler. Normally, this is not needed. See the file ‘mips.h’ for an example of this. Do not define this macro if it does not need to do anything.

**AS_NEEDS_DASH_FOR_PIPED_INPUT**

Define this macro, with no value, if the driver should give the assembler an argument consisting of a single dash, ‘-’, to instruct it to read from its standard input (which will be a pipe connected to the output of the compiler proper). This argument is given after any ‘-o’ option specifying the name of the output file.

If you do not define this macro, the assembler is assumed to read its standard input if given no non-option arguments. If your assembler cannot read standard input at all, use a '%{pipe:%e}’ construct; see ‘mips.h’ for instance.

**LINK_SPEC**

A C string constant that tells the GCC driver program options to pass to the linker. It can also specify how to translate options you give to GCC into options for GCC to pass to the linker.

Do not define this macro if it does not need to do anything.

**LIB_SPEC**

Another C string constant used much like LINK_SPEC. The difference between the two is that LIB_SPEC is used at the end of the command given to the linker.

If this macro is not defined, a default is provided that loads the standard C library from the usual place. See ‘gcc.c’.

**LIBGCC_SPEC**

Another C string constant that tells the GCC driver program how and when to place a reference to ‘libgcc.a’ into the linker command line. This constant is placed both before and after the value of LIB_SPEC.

If this macro is not defined, the GCC driver provides a default that passes the string ‘-lgcc’ to the linker.

**REAL_LIBGCC_SPEC**

By default, if ENABLE_SHARED_LIBGCC is defined, the LIBGCC_SPEC is not directly used by the driver program but is instead modified to refer to different versions of ‘libgcc.a’ depending on the values of the command line flags ‘-static’, ‘-shared’, ‘-static-libgcc’, and ‘-shared-libgcc’. On targets where these modifications are inappropriate, define REAL_LIBGCC_SPEC instead. REAL_LIBGCC_SPEC tells the driver how to place a reference to ‘libgcc’ on the link command line, but, unlike LIBGCC_SPEC, it is used unmodified.

**USE_LD_AS_NEEDED**

A macro that controls the modifications to LIBGCC_SPEC mentioned in REAL_LIBGCC_SPEC. If nonzero, a spec will be generated that uses –as-needed and the shared libgcc in place of the static exception handler library, when linking without any of ‘-static’, ‘-static-libgcc’, or ‘-shared-libgcc’.
LINK_EH_SPEC

If defined, this C string constant is added to LINK_SPEC. When USE_LD_AS_NEEDED is zero or undefined, it also affects the modifications to LIBGCC_SPEC mentioned in REAL_LIBGCC_SPEC.

STARTFILE_SPEC

Another C string constant used much like LINK_SPEC. The difference between the two is that STARTFILE_SPEC is used at the very beginning of the command given to the linker.

If this macro is not defined, a default is provided that loads the standard C startup file from the usual place. See ‘gcc.c’.

ENDFILE_SPEC

Another C string constant used much like LINK_SPEC. The difference between the two is that ENDFILE_SPEC is used at the very end of the command given to the linker.

Do not define this macro if it does not need to do anything.

THREAD_MODEL_SPEC

GCC -v will print the thread model GCC was configured to use. However, this doesn’t work on platforms that are multilibbed on thread models, such as AIX 4.3. On such platforms, define THREAD_MODEL_SPEC such that it evaluates to a string without blanks that names one of the recognized thread models. %*, the default value of this macro, will expand to the value of thread_file set in ‘config.gcc’.

SYSROOT_SUFFIX_SPEC

Define this macro to add a suffix to the target sysroot when GCC is configured with a sysroot. This will cause GCC to search for usr/lib, et al, within sysroot+suffix.

SYSROOT_HEADERS_SUFFIX_SPEC

Define this macro to add a headers suffix to the target sysroot when GCC is configured with a sysroot. This will cause GCC to pass the updated sysroot+headers_suffix to CPP, causing it to search for usr/include, et al, within sysroot+headers_suffix.

EXTRA_SPECS

Define this macro to provide additional specifications to put in the ‘specs’ file that can be used in various specifications like CC1_SPEC.

The definition should be an initializer for an array of structures, containing a string constant, that defines the specification name, and a string constant that provides the specification.

Do not define this macro if it does not need to do anything.

EXTRA_SPECS is useful when an architecture contains several related targets, which have various ..._SPECS which are similar to each other, and the maintainer would like one central place to keep these definitions.

For example, the PowerPC System V.4 targets use EXTRA_SPECS to define either _CALL_SYSV when the System V calling sequence is used or _CALL_AIX when the older AIX-based calling sequence is used.

The ‘config/rs6000/rs6000.h’ target file defines:
#define EXTRA_SPECS \
{ "cpp_sysv_default", CPP_SYSV_DEFAULT },
#define CPP_SYSV_DEFAULT ""

The ‘config/rs6000/sysv.h’ target file defines:
#undef CPP_SPEC 
#define CPP_SPEC \
"%{posix: -D_POSIX_SOURCE } \ 
%{mcall-sysv: -D_CALL_SYSV } \ 
%{!mcall-sysv: %{cpp_sysv_default} } \ 
%{msoft-float: -D_SOFT_FLOAT} %{mcpu=403: -D_SOFT_FLOAT}"
#undef CPP_SYSV_DEFAULT 
#define CPP_SYSV_DEFAULT "-D_CALL_SYSV"

while the ‘config/rs6000/eabiaix.h’ target file defines CPP_SYSV_DEFAULT as:
#undef CPP_SYSV_DEFAULT 
#define CPP_SYSV_DEFAULT "-D_CALL_AIX"

LINK_LIBGCC_SPECIAL_1 [Macro]
Define this macro if the driver program should find the library ‘libgcc.a’. If you do not define this macro, the driver program will pass the argument ‘-lgcc’ to tell the linker to do the search.

LINK_GCC_C_SEQUENCE_SPEC [Macro]
The sequence in which libgcc and libc are specified to the linker. By default this is %G %L %G.

LINK_COMMAND_SPEC [Macro]
A C string constant giving the complete command line need to execute the linker. When you do this, you will need to update your port each time a change is made to the link command line within ‘gcc.c’. Therefore, define this macro only if you need to completely redefine the command line for invoking the linker and there is no other way to accomplish the effect you need. Overriding this macro may be avoidable by overriding LINK_GCC_C_SEQUENCE_SPEC instead.

LINK_ELIMINATE_DUPLICATE_LDIRECTORIES [Macro]
A nonzero value causes collect2 to remove duplicate ‘-Ldirectory’ search directories from linking commands. Do not give it a nonzero value if removing duplicate search directories changes the linker’s semantics.

MULTILIB_DEFAULTS [Macro]
Define this macro as a C expression for the initializer of an array of string to tell the driver program which options are defaults for this target and thus do not need to be handled specially when using MULTILIB_OPTIONS.

Do not define this macro if MULTILIB_OPTIONS is not defined in the target makefile fragment or if none of the options listed in MULTILIB_OPTIONS are set by default. See Section 19.1 [Target Fragment], page 563.

RELATIVE_PREFIX_NOT_LINKDIR [Macro]
Define this macro to tell gcc that it should only translate a ‘-B’ prefix into a ‘-L’ linker option if the prefix indicates an absolute file name.
MD_EXEC_PREFIX
If defined, this macro is an additional prefix to try after STANDARD_EXEC_PREFIX. MD_EXEC_PREFIX is not searched when the compiler is built as a cross compiler. If you define MD_EXEC_PREFIX, then be sure to add it to the list of directories used to find the assembler in ‘configure.in’.

STANDARD_STARTFILE_PREFIX
Define this macro as a C string constant if you wish to override the standard choice of libdir as the default prefix to try when searching for startup files such as ‘crt0.o’. STANDARD_STARTFILE_PREFIX is not searched when the compiler is built as a cross compiler.

STANDARD_STARTFILE_PREFIX_1
Define this macro as a C string constant if you wish to override the standard choice of /lib as a prefix to try after the default prefix when searching for startup files such as ‘crt0.o’. STANDARD_STARTFILE_PREFIX_1 is not searched when the compiler is built as a cross compiler.

STANDARD_STARTFILE_PREFIX_2
Define this macro as a C string constant if you wish to override the standard choice of /lib as yet another prefix to try after the default prefix when searching for startup files such as ‘crt0.o’. STANDARD_STARTFILE_PREFIX_2 is not searched when the compiler is built as a cross compiler.

MD_STARTFILE_PREFIX
If defined, this macro supplies an additional prefix to try after the standard prefixes. MD_EXEC_PREFIX is not searched when the compiler is built as a cross compiler.

MD_STARTFILE_PREFIX_1
If defined, this macro supplies yet another prefix to try after the standard prefixes. It is not searched when the compiler is built as a cross compiler.

INIT_ENVIRONMENT
Define this macro as a C string constant if you wish to set environment variables for programs called by the driver, such as the assembler and loader. The driver passes the value of this macro to putenv to initialize the necessary environment variables.

LOCAL_INCLUDE_DIR
Define this macro as a C string constant if you wish to override the standard choice of ‘/usr/local/include’ as the default prefix to try when searching for local header files. LOCAL_INCLUDE_DIR comes before SYSTEM_INCLUDE_DIR in the search order. Cross compilers do not search either ‘/usr/local/include’ or its replacement.

SYSTEM_INCLUDE_DIR
Define this macro as a C string constant if you wish to specify a system-specific directory to search for header files before the standard directory. SYSTEM_INCLUDE_DIR comes before STANDARD_INCLUDE_DIR in the search order. Cross compilers do not use this macro and do not search the directory specified.
STANDARD_INCLUDE_DIR  [Macro]
Define this macro as a C string constant if you wish to override the standard choice of '/usr/include' as the default prefix to try when searching for header files.

Cross compilers ignore this macro and do not search either '/usr/include' or its replacement.

STANDARD_INCLUDE_COMPONENT  [Macro]
The "component" corresponding to STANDARD_INCLUDE_DIR. See INCLUDE_DEFAULTS, below, for the description of components. If you do not define this macro, no component is used.

INCLUDE_DEFAULTS  [Macro]
Define this macro if you wish to override the entire default search path for include files. For a native compiler, the default search path usually consists of GCC_INCLUDE_DIR, LOCAL_INCLUDE_DIR, SYSTEM_INCLUDE_DIR, GPLPLUS_INCLUDE_DIR, and STANDARD_INCLUDE_DIR. In addition, GPLPLUS_INCLUDE_DIR and GCC_INCLUDE_DIR are defined automatically by 'Makefile', and specify private search areas for GCC. The directory GPLPLUS_INCLUDE_DIR is used only for C++ programs.

The definition should be an initializer for an array of structures. Each array element should have four elements: the directory name (a string constant), the component name (also a string constant), a flag for C++-only directories, and a flag showing that the includes in the directory don't need to be wrapped in extern 'C' when compiling C++. Mark the end of the array with a null element.

The component name denotes what GNU package the include file is part of, if any, in all uppercase letters. For example, it might be 'GCC' or 'BINUTILS'. If the package is part of a vendor-supplied operating system, code the component name as '0'.

For example, here is the definition used for VAX/VMS:

```c
#define INCLUDE_DEFAULTS
{
    { "GNU_GXX_INCLUDE:", "G++", 1, 1},
    { "GNU_CC_INCLUDE:", "GCC", 0, 0},
    { "SYS$SYSROOT:[SYSLIB.]", 0, 0, 0},
    { ".", 0, 0, 0},
    { 0, 0, 0, 0} 
}
```

Here is the order of prefixes tried for exec files:

1. Any prefixes specified by the user with '-B'.
2. The environment variable GCC_EXEC_PREFIX or, if GCC_EXEC_PREFIX is not set and the compiler has not been installed in the configure-time prefix, the location in which the compiler has actually been installed.
3. The directories specified by the environment variable COMPILER_PATH.
4. The macro STANDARD_EXEC_PREFIX, if the compiler has been installed in the configured-time prefix.
5. The location '/usr/libexec/gcc/', but only if this is a native compiler.
6. The location '/usr/lib/gcc/', but only if this is a native compiler.
7. The macro `MD_EXEC_PREFIX`, if defined, but only if this is a native compiler.

Here is the order of prefixes tried for startfiles:
1. Any prefixes specified by the user with `-B`.
2. The environment variable `GCC_EXEC_PREFIX` or its automatically determined value based on the installed toolchain location.
3. The directories specified by the environment variable `LIBRARY_PATH` (or port-specific name; native only, cross compilers do not use this).
4. The macro `STANDARD_EXEC_PREFIX`, but only if the toolchain is installed in the configured prefix or this is a native compiler.
5. The location `/usr/lib/gcc/`, but only if this is a native compiler.
6. The macro `MD_EXEC_PREFIX`, if defined, but only if this is a native compiler.
7. The macro `MD_STARTFILE_PREFIX`, if defined, but only if this is a native compiler, or we have a target system root.
8. The macro `MD_STARTFILE_PREFIX_1`, if defined, but only if this is a native compiler, or we have a target system root.
9. The macro `STANDARD_STARTFILE_PREFIX`, with any sysroot modifications. If this path is relative it will be prefixed by `GCC_EXEC_PREFIX` and the machine suffix or `STANDARD_EXEC_PREFIX` and the machine suffix.
10. The macro `STANDARD_STARTFILE_PREFIX_1`, but only if this is a native compiler, or we have a target system root. The default for this macro is `/lib/`.
11. The macro `STANDARD_STARTFILE_PREFIX_2`, but only if this is a native compiler, or we have a target system root. The default for this macro is `/usr/lib/`.

17.3 Run-time Target Specification

Here are run-time target specifications.

`TARGET_CPU_CPP_BUILTINS ()` [Macro]

This function-like macro expands to a block of code that defines built-in preprocessor macros and assertions for the target CPU, using the functions `builtin_define`, `builtin_define_std` and `builtin_assert`. When the front end calls this macro it provides a trailing semicolon, and since it has finished command line option processing your code can use those results freely.

`builtin_assert` takes a string in the form you pass to the command-line option `-A`, such as `cpu=mips`, and creates the assertion. `builtin_define` takes a string in the form accepted by option `-D` and unconditionally defines the macro.

`builtin_define_std` takes a string representing the name of an object-like macro. If it doesn’t lie in the user’s namespace, `builtin_define_std` defines it unconditionally. Otherwise, it defines a version with two leading underscores, and another version with two leading and trailing underscores, and defines the original only if an ISO standard was not requested on the command line. For example, passing `unix` defines `__unix`, `__unix__` and possibly `unix`; passing `mips` defines `__mips`, `__mips__` and possibly `__mips`, and passing `_ABI64` defines only `_ABI64`.

You can also test for the C dialect being compiled. The variable `c_language` is set to one of `clk_c`, `clk_cplusplus` or `clk_objective_c`. Note that if we are preprocessing
assembler, this variable will be `clk_c` but the function-like macro `preprocessing_asm_p()` will return true, so you might want to check for that first. If you need to check for strict ANSI, the variable `flag_iso` can be used. The function-like macro `preprocessing_trad_p()` can be used to check for traditional preprocessing.

**TARGET_OS_CPP_BUILTINS ()**  
[Macro]  
Similarly to `TARGET_CPU_CPP_BUILTINS` but this macro is optional and is used for the target operating system instead.

**TARGET_OBJFMT_CPP_BUILTINS ()**  
[Macro]  
Similarly to `TARGET_CPU_CPP_BUILTINS` but this macro is optional and is used for the target object format. ‘elfos.h’ uses this macro to define `__ELF__`, so you probably do not need to define it yourself.

**extern int target_flags**  
[Variable]  
This variable is declared in ‘options.h’, which is included before any target-specific headers.

**int TARGET_DEFAULT_TARGET_FLAGS**  
[Target Hook]  
This variable specifies the initial value of `target_flags`. Its default setting is 0.

**bool TARGET_HANDLE_OPTION (size_t code, const char *arg, int value)**  
[Target Hook]  
This hook is called whenever the user specifies one of the target-specific options described by the ‘.opt’ definition files (see Chapter 8 [Options], page 103). It has the opportunity to do some option-specific processing and should return true if the option is valid. The default definition does nothing but return true.

`code` specifies the `OPT_name` enumeration value associated with the selected option; `name` is just a rendering of the option name in which non-alphanumeric characters are replaced by underscores. `arg` specifies the string argument and is null if no argument was given. If the option is flagged as a `UInteger` (see Section 8.2 [Option properties], page 105), `value` is the numeric value of the argument. Otherwise `value` is 1 if the positive form of the option was used and 0 if the “no-” form was.

**bool TARGET_HANDLE_C_OPTION (size_t code, const char *arg, int value)**  
[Target Hook]  
This target hook is called whenever the user specifies one of the target-specific C language family options described by the ‘.opt’ definition files (see Chapter 8 [Options], page 103). It has the opportunity to do some option-specific processing and should return true if the option is valid. The arguments are like for `TARGET_HANDLE_OPTION`. The default definition does nothing but return false.

In general, you should use `TARGET_HANDLE_OPTION` to handle options. However, if processing an option requires routines that are only available in the C (and related language) front ends, then you should use `TARGET_HANDLE_C_OPTION` instead.

**tree TARGET_OBJC_CONSTRUCT_STRING_OBJECT (tree string)**  
[Target Hook]  
Targets may provide a string object type that can be used within and between C, C++ and their respective Objective-C dialects. A string object might, for example, embed encoding and length information. These objects are considered opaque to the compiler.
and handled as references. An ideal implementation makes the composition of the string object match that of the Objective-C NSString (NXString for GNUStep), allowing efficient interworking between C-only and Objective-C code. If a target implements string objects then this hook should return a reference to such an object constructed from the normal ‘C’ string representation provided in string. At present, the hook is used by Objective-C only, to obtain a common-format string object when the target provides one.

bool TARGET_STRING_OBJECT_REF_TYPE_P(const_tree stringref) [Target Hook]
If a target implements string objects then this hook should return true if stringref is a valid reference to such an object.

void TARGET_CHECK_STRING_OBJECT_FORMAT_ARG(tree format_arg, tree args_list) [Target Hook]
If a target implements string objects then this hook should should provide a facility to check the function arguments in args_list against the format specifiers in format_arg where the type of format_arg is one recognized as a valid string reference type.

TARGET_VERSION [Macro]
This macro is a C statement to print on stderr a string describing the particular machine description choice. Every machine description should define TARGET_VERSION. For example:

```c
#ifdef MOTOROLA
#define TARGET_VERSION \ 
fprintf (stderr, " (68k, Motorola syntax)");
#else
#define TARGET_VERSION \ 
fprintf (stderr, " (68k, MIT syntax)");
#endif
```

void TARGET_OVERRIDE_OPTIONS_AFTER_CHANGE (void) [Target Hook]
This target function is similar to the hook TARGET_OPTION_OVERRIDE but is called when the optimize level is changed via an attribute or pragma or when it is reset at the end of the code affected by the attribute or pragma. It is not called at the beginning of compilation when TARGET_OPTION_OVERRIDE is called so if you want to perform these actions then, you should have TARGET_OPTION_OVERRIDE call TARGET_OVERRIDE_OPTIONS_AFTER_CHANGE.

C_COMMON_OVERRIDE_OPTIONS [Macro]
This is similar to the TARGET_OPTION_OVERRIDE hook but is only used in the C language frontends (C, Objective-C, C++, Objective-C++) and so can be used to alter option flag variables which only exist in those frontends.

const struct default_options * TARGET_OPTION_OPTIMIZATION_TABLE [Target Hook]
Some machines may desire to change what optimizations are performed for various optimization levels. This variable, if defined, describes options to enable at particular sets of optimization levels. These options are processed once just after the optimization level is determined and before the remainder of the command options have been parsed, so may be overridden by other options passed explicitly.
This processing is run once at program startup and when the optimization options are changed via `#pragma GCC optimize` or by using the `optimize` attribute.

```c
void TARGET_OPTION_INIT_STRUCT (struct gcc_options *opts) [Target Hook]
Set target-dependent initial values of fields in `opts`.
```

```c
void TARGET_OPTION_DEFAULT_PARAMS (void) [Target Hook]
Set target-dependent default values for `--param` settings, using calls to `set_default_param_value`.
```

```c
void TARGET_HELP (void) [Target Hook]
This hook is called in response to the user invoking `--target-help` on the command line. It gives the target a chance to display extra information on the target specific command line options found in its `.opt` file.
```

```c
SWITCHABLE_TARGET [Macro]
Some targets need to switch between substantially different subtargets during compilation. For example, the MIPS target has one subtarget for the traditional MIPS architecture and another for MIPS16. Source code can switch between these two subarchitectures using the `mips16` and `nomips16` attributes.

Such subtargets can differ in things like the set of available registers, the set of available instructions, the costs of various operations, and so on. GCC caches a lot of this type of information in global variables, and recomputing them for each subtarget takes a significant amount of time. The compiler therefore provides a facility for maintaining several versions of the global variables and quickly switching between them; see `target-globals.h` for details.

Define this macro to 1 if your target needs this facility. The default is 0.
```

### 17.4 Defining data structures for per-function information.

If the target needs to store information on a per-function basis, GCC provides a macro and a couple of variables to allow this. Note, just using statics to store the information is a bad idea, since GCC supports nested functions, so you can be halfway through encoding one function when another one comes along.

GCC defines a data structure called `struct function` which contains all of the data specific to an individual function. This structure contains a field called `machine` whose type is `struct machine_function *`, which can be used by targets to point to their own specific data.

If a target needs per-function specific data it should define the type `struct machine_function` and also the macro `INIT_EXPANDERS`. This macro should be used to initialize the function pointer `init_machine_status`. This pointer is explained below.

One typical use of per-function, target specific data is to create an RTX to hold the register containing the function’s return address. This RTX can then be used to implement the `__builtin_return_address` function, for level 0.

Note—earlier implementations of GCC used a single data area to hold all of the per-function information. Thus when processing of a nested function began the old per-function data had to be pushed onto a stack, and when the processing was finished, it had to be
popped off the stack. GCC used to provide function pointers called `save_machine_status` and `restore_machine_status` to handle the saving and restoring of the target specific information. Since the single data area approach is no longer used, these pointers are no longer supported.

`INIT_EXPANDERS` [Macro]
Macro called to initialize any target specific information. This macro is called once per function, before generation of any RTL has begun. The intention of this macro is to allow the initialization of the function pointer `init_machine_status`.

`void (*)(struct function *) init_machine_status` [Variable]
If this function pointer is non-NULL it will be called once per function, before function compilation starts, in order to allow the target to perform any target specific initialization of the `struct function` structure. It is intended that this would be used to initialize the `machine` of that structure.

`struct machine_function` structures are expected to be freed by GC. Generally, any memory that they reference must be allocated by using GC allocation, including the structure itself.

### 17.5 Storage Layout

Note that the definitions of the macros in this table which are sizes or alignments measured in bits do not need to be constant. They can be C expressions that refer to static variables, such as the `target_flags`. See Section 17.3 [Run-time Target], page 400.

`BITS_BIG_ENDIAN` [Macro]
Define this macro to have the value 1 if the most significant bit in a byte has the lowest number; otherwise define it to have the value zero. This means that bit-field instructions count from the most significant bit. If the machine has no bit-field instructions, then this must still be defined, but it doesn’t matter which value it is defined to. This macro need not be a constant.

This macro does not affect the way structure fields are packed into bytes or words; that is controlled by `BYTES_BIG_ENDIAN`.

`BYTES_BIG_ENDIAN` [Macro]
Define this macro to have the value 1 if the most significant byte in a word has the lowest number. This macro need not be a constant.

`WORDS_BIG_ENDIAN` [Macro]
Define this macro to have the value 1 if, in a multiword object, the most significant word has the lowest number. This applies to both memory locations and registers; GCC fundamentally assumes that the order of words in memory is the same as the order in registers. This macro need not be a constant.

`FLOAT_WORDS_BIG_ENDIAN` [Macro]
Define this macro to have the value 1 if `DFmode`, `XFmode` or `TFmode` floating point numbers are stored in memory with the word containing the sign bit at the lowest address; otherwise define it to have the value 0. This macro need not be a constant.

You need not define this macro if the ordering is the same as for multi-word integers.
BITS_PER_UNIT
[Macro]
Define this macro to be the number of bits in an addressable storage unit (byte). If you do not define this macro the default is 8.

BITS_PER_WORD
[Macro]
Number of bits in a word. If you do not define this macro, the default is BITS_PER_UNIT * UNITS_PER_WORD.

MAX_BITS_PER_WORD
[Macro]
Maximum number of bits in a word. If this is undefined, the default is BITS_PER_WORD. Otherwise, it is the constant value that is the largest value that BITS_PER_WORD can have at run-time.

UNITS_PER_WORD
[Macro]
Number of storage units in a word; normally the size of a general-purpose register, a power of two from 1 or 8.

MIN_UNITS_PER_WORD
[Macro]
Minimum number of units in a word. If this is undefined, the default is UNITS_PER_WORD. Otherwise, it is the constant value that is the smallest value that UNITS_PER_WORD can have at run-time.

POINTER_SIZE
[Macro]
Width of a pointer, in bits. You must specify a value no wider than the width of Pmode. If it is not equal to the width of Pmode, you must define POINTERS_EXTEND_UNSIGNED. If you do not specify a value the default is BITS_PER_WORD.

POINTERS_EXTEND_UNSIGNED
[Macro]
A C expression that determines how pointers should be extended from ptr_mode to either Pmode or word_mode. It is greater than zero if pointers should be zero-extended, zero if they should be sign-extended, and negative if some other sort of conversion is needed. In the last case, the extension is done by the target’s ptr_extend instruction.

You need not define this macro if the ptr_mode, Pmode and word_mode are all the same width.

PROMOTE_MODE (m, unsignedp, type)
[Macro]
A macro to update m and unsignedp when an object whose type is type and which has the specified mode and signedness is to be stored in a register. This macro is only called when type is a scalar type.

On most RISC machines, which only have operations that operate on a full register, define this macro to set m to word_mode if m is an integer mode narrower than BITS_PER_WORD. In most cases, only integer modes should be widened because wider-precision floating-point operations are usually more expensive than their narrower counterparts.

For most machines, the macro definition does not change unsignedp. However, some machines, have instructions that preferentially handle either signed or unsigned quantities of certain modes. For example, on the DEC Alpha, 32-bit loads from memory and 32-bit add instructions sign-extend the result to 64 bits. On such machines, set unsignedp according to which kind of extension is more efficient.

Do not define this macro if it would never modify m.
enum machine_mode TARGET_PROMOTE_FUNCTION_MODE (const tree type, enum machine_mode mode, int *punsignedp, const tree funtype, int for_return)

Like PROMOTE_MODE, but it is applied to outgoing function arguments or function return values. The target hook should return the new mode and possibly change *punsignedp if the promotion should change signedness. This function is called only for scalar or pointer types.

for_return allows to distinguish the promotion of arguments and return values. If it is 1, a return value is being promoted and TARGET_FUNCTION_VALUE must perform the same promotions done here. If it is 2, the returned mode should be that of the register in which an incoming parameter is copied, or the outgoing result is computed; then the hook should return the same mode as promote_mode, though the signedness may be different.

The default is to not promote arguments and return values. You can also define the hook to default_promote_function_mode_always_promote if you would like to apply the same rules given by PROMOTE_MODE.

PARM_BOUNDARY
Normal alignment required for function parameters on the stack, in bits. All stack parameters receive at least this much alignment regardless of data type. On most machines, this is the same as the size of an integer.

STACK_BOUNDARY
Define this macro to the minimum alignment enforced by hardware for the stack pointer on this machine. The definition is a C expression for the desired alignment (measured in bits). This value is used as a default if PREFERRED_STACK_BOUNDARY is not defined. On most machines, this should be the same as PARM_BOUNDARY.

PREFERRED_STACK_BOUNDARY
Define this macro if you wish to preserve a certain alignment for the stack pointer, greater than what the hardware enforces. The definition is a C expression for the desired alignment (measured in bits). This macro must evaluate to a value equal to or larger than STACK_BOUNDARY.

INCOMING_STACK_BOUNDARY
Define this macro if the incoming stack boundary may be different from PREFERRED_STACK_BOUNDARY. This macro must evaluate to a value equal to or larger than STACK_BOUNDARY.

FUNCTION_BOUNDARY
Alignment required for a function entry point, in bits.

BIGGEST_ALIGNMENT
Biggest alignment that any data type can require on this machine, in bits. Note that this is not the biggest alignment that is supported, just the biggest alignment that, when violated, may cause a fault.

MALLOC_ABI_ALIGNMENT
Alignment, in bits, a C conformant malloc implementation has to provide. If not defined, the default value is BITS_PER_WORD.
ATTRIBUTE_ALIGNED_VALUE [Macro]
   Alignment used by the __attribute__((aligned)) construct. If not defined, the
default value is BIGGEST_ALIGNMENT.

MINIMUM_ATOMIC_ALIGNMENT [Macro]
   If defined, the smallest alignment, in bits, that can be given to an object that
can be referenced in one operation, without disturbing any nearby object. Normally, this
is BITS_PER_UNIT, but may be larger on machines that don’t have byte or half-word
store operations.

BIGGEST_FIELD_ALIGNMENT [Macro]
   Biggest alignment that any structure or union field can require on this machine,
in bits. If defined, this overrides BIGGEST_ALIGNMENT for structure and union fields
only, unless the field alignment has been set by the __attribute__((aligned (n)))
construct.

ADJUST_FIELD_ALIGN(field, computed) [Macro]
   An expression for the alignment of a structure field field if the alignment computed
in the usual way (including applying of BIGGEST_ALIGNMENT and BIGGEST_FIELD_-
ALIGNMENT to the alignment) is computed. It overrides alignment only if the field
alignment has not been set by the __attribute__((aligned (n))) construct.

MAX_STACK_ALIGNMENT [Macro]
   Biggest stack alignment guaranteed by the backend. Use this macro to specify the
maximum alignment of a variable on stack.
   If not defined, the default value is STACK_BOUNDARY.

MAXOfFile_ALIGNMENT [Macro]
   Biggest alignment supported by the object file format of this machine. Use this macro
to limit the alignment which can be specified using the __attribute__((aligned
(n))) construct. If not defined, the default value is BIGGEST_ALIGNMENT.
   On systems that use ELF, the default (in 'config/elfos.h') is the largest supported
32-bit ELF section alignment representable on a 32-bit host e.g. ‘(((unsigned
HOST_WIDEST_INT) 1 << 28) * 8)’. On 32-bit ELF the largest supported section
alignment in bits is ‘(0x80000000 * 8)’, but this is not representable on 32-bit hosts.

DATA_ALIGNMENT(type, basic-align) [Macro]
   If defined, a C expression to compute the alignment for a variable in the static store.
   type is the data type, and basic-align is the alignment that the object would ordinarily
have. The value of this macro is used instead of that alignment to align the object.
   If this macro is not defined, then basic-align is used.

   One use of this macro is to increase alignment of medium-size data to make it all fit
in fewer cache lines. Another is to cause character arrays to be word-aligned so that
strcpy calls that copy constants to character arrays can be done inline.

CONSTANT_ALIGNMENT(constant, basic-align) [Macro]
   If defined, a C expression to compute the alignment given to a constant that is being
placed in memory. constant is the constant and basic-align is the alignment that
the object would ordinarily have. The value of this macro is used instead of that
alignment to align the object.
If this macro is not defined, then \texttt{basic-align} is used.
The typical use of this macro is to increase alignment for string constants to be word
aligned so that \texttt{strcpy} calls that copy constants can be done inline.

\textbf{LOCAL_ALIGNMENT (type, basic-align)} \hspace{1cm} \textbf{[Macro]}
If defined, a C expression to compute the alignment for a variable in the local store.
\texttt{type} is the data type, and \texttt{basic-align} is the alignment that the object would ordinarily
have. The value of this macro is used instead of that alignment to align the object.
If this macro is not defined, then \texttt{basic-align} is used.
One use of this macro is to increase alignment of medium-size data to make it all fit
in fewer cache lines.
If the value of this macro has a type, it should be an unsigned type.

\textbf{STACK_SLOT_ALIGNMENT (type, mode, basic-align)} \hspace{1cm} \textbf{[Macro]}
If defined, a C expression to compute the alignment for stack slot. \texttt{type} is the data
type, \texttt{mode} is the widest mode available, and \texttt{basic-align} is the alignment that the
slot would ordinarily have. The value of this macro is used instead of that alignment
to align the slot.
If this macro is not defined, then \texttt{basic-align} is used when \texttt{type} is \texttt{NULL}. Otherwise,
\textbf{LOCAL_ALIGNMENT} will be used.
This macro is to set alignment of stack slot to the maximum alignment of all possible
modes which the slot may have.
If the value of this macro has a type, it should be an unsigned type.

\textbf{LOCAL_DECL_ALIGNMENT (decl)} \hspace{1cm} \textbf{[Macro]}
If defined, a C expression to compute the alignment for a local variable \texttt{decl}.
If this macro is not defined, then \textbf{LOCAL_ALIGNMENT (TREE_TYPE (decl),
DECL_ALIGN (decl))} is used.
One use of this macro is to increase alignment of medium-size data to make it all fit
in fewer cache lines.
If the value of this macro has a type, it should be an unsigned type.

\textbf{MINIMUM_ALIGNMENT (exp, mode, align)} \hspace{1cm} \textbf{[Macro]}
If defined, a C expression to compute the minimum required alignment for dynamic
stack realignment purposes for \texttt{exp} (a type or decl), \texttt{mode}, assuming normal alignment
\texttt{align}.
If this macro is not defined, then \texttt{align} will be used.

\textbf{EMPTY_FIELD_BOUNDARY} \hspace{1cm} \textbf{[Macro]}
Alignment in bits to be given to a structure bit-field that follows an empty field such as\texttt{int : 0;}.
If \texttt{PCC\_BITFIELD\_TYPE\_MATTERS} is true, it overrides this macro.
**STRUCTURE_SIZE_BOUNDARY**

Number of bits which any structure or union’s size must be a multiple of. Each structure or union’s size is rounded up to a multiple of this.

If you do not define this macro, the default is the same as **BITS_PER_UNIT**.

**STRONG_ALIGNMENT**

Define this macro to be the value 1 if instructions will fail to work if given data not on the nominal alignment. If instructions will merely go slower in that case, define this macro as 0.

**PCC_BITFIELD_TYPE_MATTERS**

Define this if you wish to imitate the way many other C compilers handle alignment of bit-fields and the structures that contain them.

The behavior is that the type written for a named bit-field (int, short, or other integer type) imposes an alignment for the entire structure, as if the structure really did contain an ordinary field of that type. In addition, the bit-field is placed within the structure so that it would fit within such a field, not crossing a boundary for it.

Thus, on most machines, a named bit-field whose type is written as int would not cross a four-byte boundary, and would force four-byte alignment for the whole structure. (The alignment used may not be four bytes; it is controlled by the other alignment parameters.)

An unnamed bit-field will not affect the alignment of the containing structure.

If the macro is defined, its definition should be a C expression; a nonzero value for the expression enables this behavior.

Note that if this macro is not defined, or its value is zero, some bit-fields may cross more than one alignment boundary. The compiler can support such references if there are ‘insv’, ‘extv’, and ‘extzv’ insns that can directly reference memory.

The other known way of making bit-fields work is to define **STRUCTURE_SIZE_BOUNDARY** as large as **BIGGEST_ALIGNMENT**. Then every structure can be accessed with fullwords.

Unless the machine has bit-field instructions or you define **STRUCTURE_SIZE_BOUNDARY** that way, you must define **PCC_BITFIELD_TYPE_MATTERS** to have a nonzero value.

If your aim is to make GCC use the same conventions for laying out bit-fields as are used by another compiler, here is how to investigate what the other compiler does.

Compile and run this program:

```c
struct foo1
{
    char x;
    char :0;
    char y;
};

struct foo2
{
    char x;
    int :0;
    char y;
};
```
main ()
{
    printf ("Size of foo1 is %d\n",
            sizeof (struct foo1));
    printf ("Size of foo2 is %d\n",
            sizeof (struct foo2));
    exit (0);
}

If this prints 2 and 5, then the compiler's behavior is what you would get from PCC_BITFIELD_TYPE_MATTERS.

BITFIELD_NBYTES_LIMITED [Macro]
Like PCC_BITFIELD_TYPE_MATTERS except that its effect is limited to aligning a bit-field within the structure.

bool TARGET_ALIGN_ANON_BITFIELD (void) [Target Hook]
When PCC_BITFIELD_TYPE_MATTERS is true this hook will determine whether unnamed bitfields affect the alignment of the containing structure. The hook should return true if the structure should inherit the alignment requirements of an unnamed bitfield's type.

bool TARGET_NARROW_VOLATILE_BITFIELD (void) [Target Hook]
This target hook should return true if accesses to volatile bitfields should use the narrowest mode possible. It should return false if these accesses should use the bitfield container type.

The default is !TARGET.Strict-align.

MEMBER_TYPE_FORCES_BLK (field, mode) [Macro]
Return 1 if a structure or array containing field should be accessed using BLKMODE.
If field is the only field in the structure, mode is its mode, otherwise mode is VOID-mode. mode is provided in the case where structures of one field would require the structure's mode to retain the field's mode.
Normally, this is not needed.

ROUND_TYPE_ALIGN (type, computed, specified) [Macro]
Define this macro as an expression for the alignment of a type (given by type as a tree node) if the alignment computed in the usual way is computed and the alignment explicitly specified was specified.
The default is to use specified if it is larger; otherwise, use the smaller of computed and BIGGEST_ALIGNMENT

MAX_FIXED_MODE_SIZE [Macro]
An integer expression for the size in bits of the largest integer machine mode that should actually be used. All integer machine modes of this size or smaller can be used for structures and unions with the appropriate sizes. If this macro is undefined, GET_MODE_BITSIZE (DImode) is assumed.
STACK_SAVEAREA_MODE (save_level) [Macro]
If defined, an expression of type enum machine_mode that specifies the mode of the save area operand of a save_stack_level named pattern (see Section 16.9 [Standard Names], page 332). save_level is one of SAVE_BLOCK, SAVE_FUNCTION, or SAVE_NONLOCAL and selects which of the three named patterns is having its mode specified.
You need not define this macro if it always returns Pmode. You would most commonly define this macro if the save_stack_level patterns need to support both a 32- and a 64-bit mode.

STACK_SIZE_MODE [Macro]
If defined, an expression of type enum machine_mode that specifies the mode of the size increment operand of an allocate_stack named pattern (see Section 16.9 [Standard Names], page 332).
You need not define this macro if it always returns word_mode. You would most commonly define this macro if the allocate_stack pattern needs to support both a 32- and a 64-bit mode.

enum machine_mode TARGET_LIBGCC_CMP_RETURN_MODE (void) [Target Hook]
This target hook should return the mode to be used for the return value of compare instructions expanded to libgcc calls. If not defined word_mode is returned which is the right choice for a majority of targets.

enum machine_mode TARGET_LIBGCC_SHIFT_COUNT_MODE (void) [Target Hook]
This target hook should return the mode to be used for the shift count operand of shift instructions expanded to libgcc calls. If not defined word_mode is returned which is the right choice for a majority of targets.

enum machine_mode TARGET_UNWIND_WORD_MODE (void) [Target Hook]
Return machine mode to be used for _Unwind_Word type. The default is to use word_mode.

ROUND_TOWARDS_ZERO [Macro]
If defined, this macro should be true if the prevailing rounding mode is towards zero.
Defining this macro only affects the way ‘libgcc.a’ emulates floating-point arithmetic.
Not defining this macro is equivalent to returning zero.

LARGEST_EXPONENT_IS_NORMAL (size) [Macro]
This macro should return true if floats with size bits do not have a NaN or infinity representation, but use the largest exponent for normal numbers instead.
Defining this macro only affects the way ‘libgcc.a’ emulates floating-point arithmetic.
The default definition of this macro returns false for all sizes.

bool TARGET_MS_BITFIELD_LAYOUT_P (const tree record_type) [Target Hook]
This target hook returns true if bit-fields in the given record_type are to be laid out following the rules of Microsoft Visual C/C++, namely: (i) a bit-field won’t share the same storage unit with the previous bit-field if their underlying types have different
sizes, and the bit-field will be aligned to the highest alignment of the underlying types of itself and of the previous bit-field; (ii) a zero-sized bit-field will affect the alignment of the whole enclosing structure, even if it is unnamed; except that (iii) a zero-sized bit-field will be disregarded unless it follows another bit-field of nonzero size. If this hook returns true, other macros that control bit-field layout are ignored.

When a bit-field is inserted into a packed record, the whole size of the underlying type is used by one or more same-size adjacent bit-fields (that is, if its long:3, 32 bits is used in the record, and any additional adjacent long bit-fields are packed into the same chunk of 32 bits. However, if the size changes, a new field of that size is allocated). In an unpacked record, this is the same as using alignment, but not equivalent when packing.

If both MS bit-fields and ‘__attribute__((packed))’ are used, the latter will take precedence. If ‘__attribute__((packed))’ is used on a single field when MS bit-fields are in use, it will take precedence for that field, but the alignment of the rest of the structure may affect its placement.

---

### Target Hook

**bool TARGET_DECIMAL_FLOAT_SUPPORTED_P (void)**

Returns true if the target supports decimal floating point.

**bool TARGET_FIXED_POINT_SUPPORTED_P (void)**

Returns true if the target supports fixed-point arithmetic.

**void TARGET_EXPAND_TO_RTL_HOOK (void)**

This hook is called just before expansion into rtl, allowing the target to perform additional initializations or analysis before the expansion. For example, the rs6000 port uses it to allocate a scratch stack slot for use in copying SDmode values between memory and floating point registers whenever the function being expanded has any SDmode usage.

**void TARGET_INSTANTIATE_DECLS (void)**

This hook allows the backend to perform additional instantiations on rtl that are not actually in any insns yet, but will be later.

**const char * TARGET_MANGLE_TYPE (const_tree type)**

If your target defines any fundamental types, or any types your target uses should be mangled differently from the default, define this hook to return the appropriate encoding for these types as part of a C++ mangled name. The type argument is the tree structure representing the type to be mangled. The hook may be applied to trees which are not target-specific fundamental types; it should return NULL for all such types, as well as arguments it does not recognize. If the return value is not NULL, it must point to a statically-allocated string constant.

Target-specific fundamental types might be new fundamental types or qualified versions of ordinary fundamental types. Encode new fundamental types as ‘u n name’, where name is the name used for the type in source code, and n is the length of name in decimal. Encode qualified versions of ordinary types as ‘U n name code’, where name is the name used for the type qualifier in source code, n is the length of name as above, and code is the code used to represent the unqualified version of this type.
(See \texttt{write_builtin_type} in \texttt{cp/mangle.c} for the list of codes.) In both cases the spaces are for clarity; do not include any spaces in your string.

This hook is applied to types prior to typedef resolution. If the mangled name for a particular type depends only on that type’s main variant, you can perform typedef resolution yourself using \texttt{TYPE\_MAIN\_VARIANT} before mangling.

The default version of this hook always returns \texttt{NULL}, which is appropriate for a target that does not define any new fundamental types.

### 17.6 Layout of Source Language Data Types

These macros define the sizes and other characteristics of the standard basic data types used in programs being compiled. Unlike the macros in the previous section, these apply to specific features of C and related languages, rather than to fundamental aspects of storage layout.

- **\texttt{INT\_TYPE\_SIZE}**
  
  A C expression for the size in bits of the type \texttt{int} on the target machine. If you don’t define this, the default is one word.

- **\texttt{SHORT\_TYPE\_SIZE}**
  
  A C expression for the size in bits of the type \texttt{short} on the target machine. If you don’t define this, the default is half a word. (If this would be less than one storage unit, it is rounded up to one unit.)

- **\texttt{LONG\_TYPE\_SIZE}**
  
  A C expression for the size in bits of the type \texttt{long} on the target machine. If you don’t define this, the default is one word.

- **\texttt{ADA\_LONG\_TYPE\_SIZE}**
  
  On some machines, the size used for the Ada equivalent of the type \texttt{long} by a native Ada compiler differs from that used by C. In that situation, define this macro to be a C expression to be used for the size of that type. If you don’t define this, the default is the value of \texttt{LONG\_TYPE\_SIZE}.

- **\texttt{LONG\_LONG\_TYPE\_SIZE}**
  
  A C expression for the size in bits of the type \texttt{long long} on the target machine. If you don’t define this, the default is two words. If you want to support GNU Ada on your machine, the value of this macro must be at least 64.

- **\texttt{CHAR\_TYPE\_SIZE}**
  
  A C expression for the size in bits of the type \texttt{char} on the target machine. If you don’t define this, the default is \texttt{BITS\_PER\_UNIT}.

- **\texttt{BOOL\_TYPE\_SIZE}**
  
  A C expression for the size in bits of the C++ type \texttt{bool} and C99 type \texttt{_Bool} on the target machine. If you don’t define this, and you probably shouldn’t, the default is \texttt{CHAR\_TYPE\_SIZE}.

- **\texttt{FLOAT\_TYPE\_SIZE}**
  
  A C expression for the size in bits of the type \texttt{float} on the target machine. If you don’t define this, the default is one word.
DOUBLE_TYPE_SIZE
A C expression for the size in bits of the type double on the target machine. If you don’t define this, the default is two words.

LONG_DOUBLE_TYPE_SIZE
A C expression for the size in bits of the type long double on the target machine. If you don’t define this, the default is two words.

SHORT_FRACT_TYPE_SIZE
A C expression for the size in bits of the type short Fract on the target machine. If you don’t define this, the default is BITS_PER_UNIT.

FRACT_TYPE_SIZE
A C expression for the size in bits of the type Fract on the target machine. If you don’t define this, the default is BITS_PER_UNIT * 2.

LONG_FRACT_TYPE_SIZE
A C expression for the size in bits of the type long Fract on the target machine. If you don’t define this, the default is BITS_PER_UNIT * 4.

LONG_LONG_FRACT_TYPE_SIZE
A C expression for the size in bits of the type long long Fract on the target machine. If you don’t define this, the default is BITS_PER_UNIT * 8.

SHORT_ACCUM_TYPE_SIZE
A C expression for the size in bits of the type short Accum on the target machine. If you don’t define this, the default is BITS_PER_UNIT * 2.

ACCUM_TYPE_SIZE
A C expression for the size in bits of the type Accum on the target machine. If you don’t define this, the default is BITS_PER_UNIT * 4.

LONG_ACCUM_TYPE_SIZE
A C expression for the size in bits of the type long Accum on the target machine. If you don’t define this, the default is BITS_PER_UNIT * 8.

LONG_LONG_ACCUM_TYPE_SIZE
A C expression for the size in bits of the type long long Accum on the target machine. If you don’t define this, the default is BITS_PER_UNIT * 16.

LIBGCC2_LONG_DOUBLE_TYPE_SIZE
Define this macro if LONG_DOUBLE_TYPE_SIZE is not constant or if you want routines in ‘libgcc2.a’ for a size other than LONG_DOUBLE_TYPE_SIZE. If you don’t define this, the default is LONG_DOUBLE_TYPE_SIZE.

LIBGCC2_HAS_DF_MODE
Define this macro if neither DOUBLE_TYPE_SIZE nor LIBGCC2_LONG_DOUBLE_TYPE_SIZE is DFmode but you want DFmode routines in ‘libgcc2.a’ anyway. If you don’t define this and either DOUBLE_TYPE_SIZE or LIBGCC2_LONG_DOUBLE_TYPE_SIZE is 64 then the default is 1, otherwise it is 0.
LIBGCC2_HAS_XF_MODE  [Macro]
Define this macro if LIBGCC2_LONG_DOUBLE_TYPE_SIZE is not XFmode but you want
XFmode routines in ‘libgcc2.a’ anyway. If you don’t define this and LIBGCC2_LONG_-
DOUBLE_TYPE_SIZE is 80 then the default is 1, otherwise it is 0.

LIBGCC2_HAS_TF_MODE  [Macro]
Define this macro if LIBGCC2_LONG_DOUBLE_TYPE_SIZE is not TFmode but you want
TFmode routines in ‘libgcc2.a’ anyway. If you don’t define this and LIBGCC2_LONG_-
DOUBLE_TYPE_SIZE is 128 then the default is 1, otherwise it is 0.

SF_SIZE                [Macro]
DF_SIZE                [Macro]
XF_SIZE                [Macro]
TF_SIZE                [Macro]
Define these macros to be the size in bits of the mantissa of SFmode, DFmode, XFmode
and TFmode values, if the defaults in ‘libgcc2.h’ are inappropriate. By default,
FLT_MANT_DIG is used for SF_SIZE, LDBL_MANT_DIG for XF_SIZE and TF_SIZE, and
DBL_MANT_DIG or LDBL_MANT_DIG for DF_SIZE according to whether DOUBLE_TYPE_-
SIZE or LIBGCC2_LONG_DOUBLE_TYPE_SIZE is 64.

TARGET_FLT_EVAL_METHOD  [Macro]
A C expression for the value for FLT_EVAL_METHOD in ‘float.h’, assuming, if appli-
cable, that the floating-point control word is in its default state. If you do not define
this macro the value of FLT_EVAL_METHOD will be zero.

WIDEST_HARDWARE_FP_SIZE  [Macro]
A C expression for the size in bits of the widest floating-point format supported by
the hardware. If you define this macro, you must specify a value less than or equal
to the value of LONG_DOUBLE_TYPE_SIZE. If you do not define this macro, the value
of LONG_DOUBLE_TYPE_SIZE is the default.

DEFAULT_SIGNED_CHAR  [Macro]
An expression whose value is 1 or 0, according to whether the type char should be
signed or unsigned by default. The user can always override this default with the
options ‘-fsigned-char’ and ‘-funsigned-char’.

bool TARGET_DEFAULT_SHORT_ENUMS (void)  [Target Hook]
This target hook should return true if the compiler should give an enum type only
as many bytes as it takes to represent the range of possible values of that type. It
should return false if all enum types should be allocated like int.
The default is to return false.

SIZE_TYPE               [Macro]
A C expression for a string describing the name of the data type to use for size values.
The typedef name size_t is defined using the contents of the string.
The string can contain more than one keyword. If so, separate them with spaces,
and write first any length keyword, then unsigned if appropriate, and finally int.
The string must exactly match one of the data type names defined in the function
init_decl_processing in the file ‘c-decl.c’. You may not omit int or change the order—that would cause the compiler to crash on startup.

If you don’t define this macro, the default is "long unsigned int".

PTRDIFF_TYPE  [Macro]
A C expression for a string describing the name of the data type to use for the result of subtracting two pointers. The typedef name ptrdiff_t is defined using the contents of the string. See SIZE_TYPE above for more information.

If you don’t define this macro, the default is "long int".

WCHAR_TYPE  [Macro]
A C expression for a string describing the name of the data type to use for wide characters. The typedef name wchar_t is defined using the contents of the string. See SIZE_TYPE above for more information.

If you don’t define this macro, the default is "int".

WCHAR_TYPE_SIZE  [Macro]
A C expression for the size in bits of the data type for wide characters. This is used in cpp, which cannot make use of WCHAR_TYPE.

WINT_TYPE  [Macro]
A C expression for a string describing the name of the data type to use for wide characters passed to printf and returned from getwc. The typedef name wint_t is defined using the contents of the string. See SIZE_TYPE above for more information.

If you don’t define this macro, the default is "unsigned int".

INTMAX_TYPE  [Macro]
A C expression for a string describing the name of the data type that can represent any value of any standard or extended signed integer type. The typedef name intmax_t is defined using the contents of the string. See SIZE_TYPE above for more information.

If you don’t define this macro, the default is the first of "int", "long int", or "long long int" that has as much precision as long long int.

UINTMAX_TYPE  [Macro]
A C expression for a string describing the name of the data type that can represent any value of any standard or extended unsigned integer type. The typedef name uintmax_t is defined using the contents of the string. See SIZE_TYPE above for more information.

If you don’t define this macro, the default is the first of "unsigned int", "long unsigned int", or "long long unsigned int" that has as much precision as long long unsigned int.

SIG_ATOMIC_TYPE  [Macro]
INT8_TYPE  [Macro]
INT16_TYPE  [Macro]
INT32_TYPE  [Macro]
INT64_TYPE  [Macro]
UINT8_TYPE  [Macro]
C expressions for the standard types `sig_atomic_t`, `int8_t`, `int16_t`, `int32_t`, `int64_t`, `uint8_t`, `uint16_t`, `uint32_t`, `uint64_t`, `int_least8_t`, `int_least16_t`, `int_least32_t`, `int_least64_t`, `uint_least8_t`, `uint_least16_t`, `uint_least32_t`, `uint_least64_t`, `int_fast8_t`, `int_fast16_t`, `int_fast32_t`, `int_fast64_t`, `uint_fast8_t`, `uint_fast16_t`, `uint_fast32_t`, `uint_fast64_t`, `intptr_t`, and `uintptr_t`. See `SIZE_TYPE` above for more information.

If any of these macros evaluates to a null pointer, the corresponding type is not supported; if GCC is configured to provide `<stdint.h>` in such a case, the header provided may not conform to C99, depending on the type in question. The defaults for all of these macros are null pointers.

**TARGET_PTRMEMFUNC_VBIT_LOCATION**

The C++ compiler represents a pointer-to-member-function with a struct that looks like:

```c
struct {
    union {
        void (*fn)();
        ptrdiff_t vtable_index;
    };
    ptrdiff_t delta;
};
```

The C++ compiler must use one bit to indicate whether the function that will be called through a pointer-to-member-function is virtual. Normally, we assume that the low-order bit of a function pointer must always be zero. Then, by ensuring that the vtable_index is odd, we can distinguish which variant of the union is in use. But, on some platforms function pointers can be odd, and so this doesn’t work. In that
case, we use the low-order bit of the \texttt{delta} field, and shift the remainder of the \texttt{delta} field to the left.

GCC will automatically make the right selection about where to store this bit using the \texttt{FUNCTION_BOUNDARY} setting for your platform. However, some platforms such as ARM/Thumb have \texttt{FUNCTION_BOUNDARY} set such that functions always start at even addresses, but the lowest bit of pointers to functions indicate whether the function at that address is in ARM or Thumb mode. If this is the case of your architecture, you should define this macro to \texttt{ptrmemfunc_vbit_in_delta}.

In general, you should not have to define this macro. On architectures in which function addresses are always even, according to \texttt{FUNCTION_BOUNDARY}, GCC will automatically define this macro to \texttt{ptrmemfunc_vbit_in_pfn}.

\textbf{TARGET_VTABLE_USES_DESCRIPTORS} \hfill [Macro]

Normally, the C++ compiler uses function pointers in vtables. This macro allows the target to change to use “function descriptors” instead. Function descriptors are found on targets for whom a function pointer is actually a small data structure. Normally the data structure consists of the actual code address plus a data pointer to which the function’s data is relative.

If vtables are used, the value of this macro should be the number of words that the function descriptor occupies.

\textbf{TARGET_VTABLE_ENTRY_ALIGN} \hfill [Macro]

By default, the vtable entries are void pointers, so the alignment is the same as pointer alignment. The value of this macro specifies the alignment of the vtable entry in bits. It should be defined only when special alignment is necessary. */

\textbf{TARGET_VTABLE_DATA_ENTRY_DISTANCE} \hfill [Macro]

There are a few non-descriptor entries in the vtable at offsets below zero. If these entries must be padded (say, to preserve the alignment specified by \texttt{TARGET_VTABLE_ENTRY_ALIGN}), set this to the number of words in each data entry.

17.7 Register Usage

This section explains how to describe what registers the target machine has, and how (in general) they can be used.

The description of which registers a specific instruction can use is done with register classes; see Section 17.8 [Register Classes], page 425. For information on using registers to access a stack frame, see Section 17.10.4 [Frame Registers], page 443. For passing values in registers, see Section 17.10.7 [Register Arguments], page 449. For returning values in registers, see Section 17.10.8 [Scalar Return], page 455.

17.7.1 Basic Characteristics of Registers

Registers have various characteristics.

\textbf{FIRST_PSEUDO_REGISTER} \hfill [Macro]

Number of hardware registers known to the compiler. They receive numbers 0 through \texttt{FIRST_PSEUDO_REGISTER}-1; thus, the first pseudo register’s number really is assigned the number \texttt{FIRST_PSEUDO_REGISTER}. 

**FIXED_REGISTERS**  
An initializer that says which registers are used for fixed purposes all throughout the compiled code and are therefore not available for general allocation. These would include the stack pointer, the frame pointer (except on machines where that can be used as a general register when no frame pointer is needed), the program counter on machines where that is considered one of the addressable registers, and any other numbered register with a standard use.

This information is expressed as a sequence of numbers, separated by commas and surrounded by braces. The \( n \)th number is 1 if register \( n \) is fixed, 0 otherwise.

The table initialized from this macro, and the table initialized by the following one, may be overridden at run time either automatically, by the actions of the macro `CONDITIONAL_REGISTER_USAGE`, or by the user with the command options `-ffixed-reg`, `-fcall-used-reg` and `-fcall-saved-reg`.

**CALL_USED_REGISTERS**  
Like `FIXED_REGISTERS` but has 1 for each register that is clobbered (in general) by function calls as well as for fixed registers. This macro therefore identifies the registers that are not available for general allocation of values that must live across function calls.

If a register has 0 in `CALL_USED_REGISTERS`, the compiler automatically saves it on function entry and restores it on function exit, if the register is used within the function.

**CALL_REALLY_USED_REGISTERS**  
Like `CALL_USED_REGISTERS` except this macro doesn’t require that the entire set of `FIXED_REGISTERS` be included. (`CALL_USED_REGISTERS` must be a superset of `FIXED_REGISTERS`). This macro is optional. If not specified, it defaults to the value of `CALL_USED_REGISTERS`.

**HARD_REGNO_CALL_PART_CLOBBERED**  
A C expression that is nonzero if it is not permissible to store a value of mode `mode` in hard register number `regno` across a call without some part of it being clobbered. For most machines this macro need not be defined. It is only required for machines that do not preserve the entire contents of a register across a call.

**void TARGET_CONDITIONAL_REGISTER_USAGE (void)**  
This hook may conditionally modify five variables `fixed_regs`, `call_used_regs`, `global_regs`, `reg_names`, and `reg_class_contents`, to take into account any dependence of these register sets on target flags. The first three of these are of type `char []` (interpreted as Boolean vectors). `global_regs` is a `const char *[]`, and `reg_class_contents` is a `HARD_REG_SET`. Before the macro is called, `fixed_regs`, `call_used_regs`, `reg_class_contents`, and `reg_names` have been initialized from `FIXED_REGISTERS`, `CALL_USED_REGISTERS`, `REG_CLASS_CONTENTS`, and `REGISTER_NAMES`, respectively. `global_regs` has been cleared, and any `-ffixed-reg`, `-fcall-used-reg` and `-fcall-saved-reg` command options have been applied.

If the usage of an entire class of registers depends on the target flags, you may indicate this to GCC by using this macro to modify `fixed_regs` and `call_used_regs` to 1 for
each of the registers in the classes which should not be used by GCC. Also define the macro `REG_CLASS_FROM_LETTER / REG_CLASS_FROM_CONSTRAINT` to return `NO_REGS` if it is called with a letter for a class that shouldn’t be used.

(However, if this class is not included in `GENERAL_REGS` and all of the insn patterns whose constraints permit this class are controlled by target switches, then GCC will automatically avoid using these registers when the target switches are opposed to them.)

**INCOMING_REGNO** *(out)*

Macro

Define this macro if the target machine has register windows. This C expression returns the register number as seen by the called function corresponding to the register number `out` as seen by the calling function. Return `out` if register number `out` is not an outbound register.

**OUTGOING_REGNO** *(in)*

Macro

Define this macro if the target machine has register windows. This C expression returns the register number as seen by the calling function corresponding to the register number `in` as seen by the called function. Return `in` if register number `in` is not an inbound register.

**LOCAL_REGNO** *(regno)*

Macro

Define this macro if the target machine has register windows. This C expression returns true if the register is call-saved but is in the register window. Unlike most call-saved registers, such registers need not be explicitly restored on function exit or during non-local gotos.

**PC_REGNUM**

Macro

If the program counter has a register number, define this as that register number. Otherwise, do not define it.

### 17.7.2 Order of Allocation of Registers

Registers are allocated in order.

**REG_ALLOC_ORDER**

Macro

If defined, an initializer for a vector of integers, containing the numbers of hard registers in the order in which GCC should prefer to use them (from most preferred to least).

If this macro is not defined, registers are used lowest numbered first (all else being equal).

One use of this macro is on machines where the highest numbered registers must always be saved and the save-multiple-registers instruction supports only sequences of consecutive registers. On such machines, define `REG_ALLOC_ORDER` to be an initializer that lists the highest numbered allocable register first.

**ADJUST_REG_ALLOC_ORDER**

Macro

A C statement (sans semicolon) to choose the order in which to allocate hard registers for pseudo-registers local to a basic block.
Store the desired register order in the array \texttt{reg\_alloc\_order}. Element 0 should be the register to allocate first; element 1, the next register; and so on.

The macro body should not assume anything about the contents of \texttt{reg\_alloc\_order} before execution of the macro.

On most machines, it is not necessary to define this macro.

\textbf{HONOR\_REG\_ALLOC\_ORDER} \hfill [Macro]

Normally, IRA tries to estimate the costs for saving a register in the prologue and restoring it in the epilogue. This discourages it from using call-saved registers. If a machine wants to ensure that IRA allocates registers in the order given by \texttt{REG\_ALLOC\_ORDER} even if some call-saved registers appear earlier than call-used ones, this macro should be defined.

\textbf{IRA\_HARD\_REGNO\_ADD\_COST\_MULTIPLIER} \hfill (\texttt{regno}) \hfill [Macro]

In some case register allocation order is not enough for the Integrated Register Allocator (IRA) to generate a good code. If this macro is defined, it should return a floating point value based on \texttt{regno}. The cost of using \texttt{regno} for a pseudo will be increased by approximately the pseudo's usage frequency times the value returned by this macro. Not defining this macro is equivalent to having it always return 0.0.

On most machines, it is not necessary to define this macro.

### 17.7.3 How Values Fit in Registers

This section discusses the macros that describe which kinds of values (specifically, which machine modes) each register can hold, and how many consecutive registers are needed for a given mode.

\textbf{HARD\_REGNO\_NREGS} \hfill (\texttt{regno}, \texttt{mode}) \hfill [Macro]

A C expression for the number of consecutive hard registers, starting at register number \texttt{regno}, required to hold a value of mode \texttt{mode}. This macro must never return zero, even if a register cannot hold the requested mode - indicate that with \texttt{HARD\_REGNO\_MODE\_OK} and/or \texttt{CANNOT\_CHANGE\_MODE\_CLASS} instead.

On a machine where all registers are exactly one word, a suitable definition of this macro is

```c
#define HARD_REGNO_NREGS(REGNO, MODE) \
   ((GET_MODE_SIZE(MODE) + UNITS_PER_WORD - 1) \ 
    / UNITS_PER_WORD)
```

\textbf{HARD\_REGNO\_NREGS\_HAS\_PADDING} \hfill (\texttt{regno}, \texttt{mode}) \hfill [Macro]

A C expression that is nonzero if a value of mode \texttt{mode}, stored in memory, ends with padding that causes it to take up more space than in registers starting at register number \texttt{regno} (as determined by multiplying GCC's notion of the size of the register when containing this mode by the number of registers returned by \texttt{HARD\_REGNO\_NREGS}). By default this is zero.

For example, if a floating-point value is stored in three 32-bit registers but takes up 128 bits in memory, then this would be nonzero.

This macros only needs to be defined if there are cases where \texttt{subreg\_get\_info} would otherwise wrongly determine that a \texttt{subreg} can be represented by an offset to the
register number, when in fact such a subreg would contain some of the padding not stored in registers and so not be representable.

**HARD_REGNO_NREGS_WITH_PADDING (regno, mode)**

For values of regno and mode for which **HARD_REGNO_NREGS_HAS_PADDING** returns nonzero, a C expression returning the greater number of registers required to hold the value including any padding. In the example above, the value would be four.

**REGMODE_NATURAL_SIZE (mode)**

Define this macro if the natural size of registers that hold values of mode mode is not the word size. It is a C expression that should give the natural size in bytes for the specified mode. It is used by the register allocator to try to optimize its results. This happens for example on SPARC 64-bit where the natural size of floating-point registers is still 32-bit.

**HARD_REGNO_MODE_OK (regno, mode)**

A C expression that is nonzero if it is permissible to store a value of mode mode in hard register number regno (or in several registers starting with that one). For a machine where all registers are equivalent, a suitable definition is

```
#define HARD_REGNO_MODE_OK(REGNO, MODE) 1
```

You need not include code to check for the numbers of fixed registers, because the allocation mechanism considers them to be always occupied.

On some machines, double-precision values must be kept in even/odd register pairs. You can implement that by defining this macro to reject odd register numbers for such modes.

The minimum requirement for a mode to be OK in a register is that the ‘movmode’ instruction pattern support moves between the register and other hard register in the same class and that moving a value into the register and back out not alter it.

Since the same instruction used to move word_mode will work for all narrower integer modes, it is not necessary on any machine for **HARD_REGNO_MODE_OK** to distinguish between these modes, provided you define patterns ‘movhi’, etc., to take advantage of this. This is useful because of the interaction between **HARD_REGNO_MODE_OK** and **MODES_TIEABLE_P**; it is very desirable for all integer modes to be tieable.

Many machines have special registers for floating point arithmetic. Often people assume that floating point machine modes are allowed only in floating point registers. This is not true. Any registers that can hold integers can safely hold a floating point machine mode, whether or not floating arithmetic can be done on it in those registers. Integer move instructions can be used to move the values.

On some machines, though, the converse is true: fixed-point machine modes may not go in floating registers. This is true if the floating registers normalize any value stored in them, because storing a non-floating value there would garble it. In this case, **HARD_REGNO_MODE_OK** should reject fixed-point machine modes in floating registers. But if the floating registers do not automatically normalize, if you can store any bit pattern in one and retrieve it unchanged without a trap, then any machine mode may go in a floating register, so you can define this macro to say so.

The primary significance of special floating registers is rather that they are the registers acceptable in floating point arithmetic instructions. However, this is of no
concern to HARD_REGNO_MODE_OK. You handle it by writing the proper constraints for those instructions.

On some machines, the floating registers are especially slow to access, so that it is better to store a value in a stack frame than in such a register if floating point arithmetic is not being done. As long as the floating registers are not in class GENERAL_REGS, they will not be used unless some pattern’s constraint asks for one.

HARD_REGNO_RENAME_OK (from, to) [Macro]
A C expression that is nonzero if it is OK to rename a hard register from to another hard register to.

One common use of this macro is to prevent renaming of a register to another register that is not saved by a prologue in an interrupt handler.

The default is always nonzero.

MODES_TIEABLE_P (mode1, mode2) [Macro]
A C expression that is nonzero if a value of mode mode1 is accessible in mode mode2 without copying.

If HARD_REGNO_MODE_OK (r, mode1) and HARD_REGNO_MODE_OK (r, mode2) are always the same for any r, then MODES_TIEABLE_P (mode1, mode2) should be nonzero. If they differ for any r, you should define this macro to return zero unless some other mechanism ensures the accessibility of the value in a narrower mode.

You should define this macro to return nonzero in as many cases as possible since doing so will allow GCC to perform better register allocation.

bool TARGET_HARD_REGNO_SCRATCH_OK (unsigned int regno) [Target Hook]
This target hook should return true if it is OK to use a hard register regno as scratch reg in peephole2.

One common use of this macro is to prevent using of a register that is not saved by a prologue in an interrupt handler.

The default version of this hook always returns true.

AVOID_CCMODE_COPIES [Macro]
Define this macro if the compiler should avoid copies to/from CCmode registers. You should only define this macro if support for copying to/from CCmode is incomplete.

17.7.4 Handling Leaf Functions

On some machines, a leaf function (i.e., one which makes no calls) can run more efficiently if it does not make its own register window. Often this means it is required to receive its arguments in the registers where they are passed by the caller, instead of the registers where they would normally arrive.

The special treatment for leaf functions generally applies only when other conditions are met; for example, often they may use only those registers for its own variables and temporaries. We use the term “leaf function” to mean a function that is suitable for this special handling, so that functions with no calls are not necessarily “leaf functions”.

GCC assigns register numbers before it knows whether the function is suitable for leaf function treatment. So it needs to renumber the registers in order to output a leaf function. The following macros accomplish this.
LEAF_REGISTERS [Macro]
Name of a char vector, indexed by hard register number, which contains 1 for a register that is allowable in a candidate for leaf function treatment.
If leaf function treatment involves renumbering the registers, then the registers marked here should be the ones before renumbering—those that GCC would ordinarily allocate. The registers which will actually be used in the assembler code, after renumbering, should not be marked with 1 in this vector.
Define this macro only if the target machine offers a way to optimize the treatment of leaf functions.

LEAF_REG_REMAP (regno) [Macro]
A C expression whose value is the register number to which regno should be renumbered, when a function is treated as a leaf function.
If regno is a register number which should not appear in a leaf function before renumbering, then the expression should yield -1, which will cause the compiler to abort.
Define this macro only if the target machine offers a way to optimize the treatment of leaf functions, and registers need to be renumbered to do this.

TARGET_ASM_FUNCTION_PROLOGUE and TARGET_ASM_FUNCTION_EPILOGUE must usually treat leaf functions specially. They can test the C variable current_function_is_leaf which is nonzero for leaf functions. current_function_is_leaf is set prior to local register allocation and is valid for the remaining compiler passes. They can also test the C variable current_function_uses_only_leaf_regs which is nonzero for leaf functions which only use leaf registers. current_function_uses_only_leaf_regs is valid after all passes that modify the instructions have been run and is only useful if LEAF_REGISTERS is defined.

17.7.5 Registers That Form a Stack
There are special features to handle computers where some of the “registers” form a stack. Stack registers are normally written by pushing onto the stack, and are numbered relative to the top of the stack.
Currently, GCC can only handle one group of stack-like registers, and they must be consecutively numbered. Furthermore, the existing support for stack-like registers is specific to the 80387 floating point coprocessor. If you have a new architecture that uses stack-like registers, you will need to do substantial work on ‘reg-stack.c’ and write your machine description to cooperate with it, as well as defining these macros.

STACK_REGS [Macro]
Define this if the machine has any stack-like registers.

STACK_REG_COVER_CLASS [Macro]
This is a cover class containing the stack registers. Define this if the machine has any stack-like registers.

FIRST_STACK_REG [Macro]
The number of the first stack-like register. This one is the top of the stack.

LAST_STACK_REG [Macro]
The number of the last stack-like register. This one is the bottom of the stack.
17.8 Register Classes

On many machines, the numbered registers are not all equivalent. For example, certain registers may not be allowed for indexed addressing; certain registers may not be allowed in some instructions. These machine restrictions are described to the compiler using register classes.

You define a number of register classes, giving each one a name and saying which of the registers belong to it. Then you can specify register classes that are allowed as operands to particular instruction patterns.

In general, each register will belong to several classes. In fact, one class must be named ALL_REGS and contain all the registers. Another class must be named NO_REGS and contain no registers. Often the union of two classes will be another class; however, this is not required.

One of the classes must be named GENERAL_REGS. There is nothing terribly special about the name, but the operand constraint letters ‘r’ and ‘g’ specify this class. If GENERAL_REGS is the same as ALL_REGS, just define it as a macro which expands to ALL_REGS.

Order the classes so that if class x is contained in class y then x has a lower class number than y.

The way classes other than GENERAL_REGS are specified in operand constraints is through machine-dependent operand constraint letters. You can define such letters to correspond to various classes, then use them in operand constraints.

You should define a class for the union of two classes whenever some instruction allows both classes. For example, if an instruction allows either a floating point (coprocessor) register or a general register for a certain operand, you should define a class FLOAT_OR_GENERAL_REGS which includes both of them. Otherwise you will get suboptimal code, or even internal compiler errors when reload cannot find a register in the the class computed via reg_class_subunion.

You must also specify certain redundant information about the register classes: for each class, which classes contain it and which ones are contained in it; for each pair of classes, the largest class contained in their union.

When a value occupying several consecutive registers is expected in a certain class, all the registers used must belong to that class. Therefore, register classes cannot be used to enforce a requirement for a register pair to start with an even-numbered register. The way to specify this requirement is with HARD_REGNO_MODE_OK.

Register classes used for input-operands of bitwise-and or shift instructions have a special requirement: each such class must have, for each fixed-point machine mode, a subclass whose registers can transfer that mode to or from memory. For example, on some machines, the operations for single-byte values (QImode) are limited to certain registers. When this is so, each register class that is used in a bitwise-and or shift instruction must have a subclass consisting of registers from which single-byte values can be loaded or stored. This is so that PREFERRED_RELOAD_CLASS can always have a possible value to return.

```c
enum reg_class
{
  [Data type]
  NO_REGS,
  ALL_REGS,
  //...

#define reg_class
```

An enumerated type that must be defined with all the register class names as enumerated values. NO_REGS must be first. ALL_REGS must be the last register class,
followed by one more enumerated value, \texttt{LIM\_REG\_CLASSES}, which is not a register class but rather tells how many classes there are.

Each register class has a number, which is the value of casting the class name to type \texttt{int}. The number serves as an index in many of the tables described below.

\textbf{\texttt{N\_REG\_CLASSES}} \hspace{1cm} \textbf{[Macro]}

The number of distinct register classes, defined as follows:

\begin{center}
\begin{verbatim}
#define N_REG_CLASSES (int) LIM_REG_CLASSES
\end{verbatim}
\end{center}

\textbf{\texttt{REG\_CLASS\_NAMES}} \hspace{1cm} \textbf{[Macro]}

An initializer containing the names of the register classes as C string constants. These names are used in writing some of the debugging dumps.

\textbf{\texttt{REG\_CLASS\_CONTENTS}} \hspace{1cm} \textbf{[Macro]}

An initializer containing the contents of the register classes, as integers which are bit masks. The \textit{n}th integer specifies the contents of class \textit{n}. The way the integer \textit{mask} is interpreted is that register \textit{r} is in the class if \textit{mask} \& \((1 << r)\) is 1.

When the machine has more than 32 registers, an integer does not suffice. Then the integers are replaced by sub-initializers, braced groupings containing several integers. Each sub-initializer must be suitable as an initializer for the type \texttt{HARD\_REG\_SET} which is defined in \texttt{`hard-reg-set.h'}. In this situation, the first integer in each sub-initializer corresponds to registers 0 through 31, the second integer to registers 32 through 63, and so on.

\textbf{\texttt{REGNO\_REG\_CLASS} \((\texttt{regno})\)} \hspace{1cm} \textbf{[Macro]}

A C expression whose value is a register class containing hard register \texttt{regno}. In general there is more than one such class; choose a class which is minimal, meaning that no smaller class also contains the register.

\textbf{\texttt{BASE\_REG\_CLASS}} \hspace{1cm} \textbf{[Macro]}

A macro whose definition is the name of the class to which a valid base register must belong. A base register is one used in an address which is the register value plus a displacement.

\textbf{\texttt{MODE\_BASE\_REG\_CLASS} \((\texttt{mode})\)} \hspace{1cm} \textbf{[Macro]}

This is a variation of the \texttt{BASE\_REG\_CLASS} macro which allows the selection of a base register in a mode dependent manner. If \texttt{mode} is \texttt{VOIDmode} then it should return the same value as \texttt{BASE\_REG\_CLASS}.

\textbf{\texttt{MODE\_BASE\_REG\_REG\_CLASS} \((\texttt{mode})\)} \hspace{1cm} \textbf{[Macro]}

A C expression whose value is the register class to which a valid base register must belong in order to be used in a base plus index register address. You should define this macro if base plus index addresses have different requirements than other base register uses.

\textbf{\texttt{MODE\_CODE\_BASE\_REG\_CLASS} \((\texttt{mode, outer\_code, index\_code})\)} \hspace{1cm} \textbf{[Macro]}

A C expression whose value is the register class to which a valid base register must belong. \texttt{outer\_code} and \texttt{index\_code} define the context in which the base register occurs. \texttt{outer\_code} is the code of the immediately enclosing expression (\texttt{MEM} for the
top level of an address, ADDRESS for something that occurs in an address_operand).

index_code is the code of the corresponding index expression if outer_code is PLUS; SCRATCH otherwise.

INDEX_REG_CLASS [Macro]
A macro whose definition is the name of the class to which a valid index register must belong. An index register is one used in an address where its value is either multiplied by a scale factor or added to another register (as well as added to a displacement).

REGNO_OK_FOR_BASE_P (num) [Macro]
A C expression which is nonzero if register number num is suitable for use as a base register in operand addresses.

REGNO_MODE_OK_FOR_BASE_P (num, mode) [Macro]
A C expression that is just like REGNO_OK_FOR_BASE_P, except that that expression may examine the mode of the memory reference in mode. You should define this macro if the mode of the memory reference affects whether a register may be used as a base register. If you define this macro, the compiler will use it instead of REGNO_OK_FOR_BASE_P. The mode may be VOIDmode for addresses that appear outside a MEM, i.e., as an address_operand.

REGNO_MODE_OK_FOR_REG_BASE_P (num, mode) [Macro]
A C expression which is nonzero if register number num is suitable for use as a base register in base plus index operand addresses, accessing memory in mode mode. It may be either a suitable hard register or a pseudo register that has been allocated such a hard register. You should define this macro if base plus index addresses have different requirements than other base register uses.

Use of this macro is deprecated; please use the more general REGNO_MODE_CODE_OK_FOR_BASE_P.

REGNO_MODE_CODE_OK_FOR_BASE_P (num, mode, outer_code, index_code) [Macro]
A C expression that is just like REGNO_MODE_OK_FOR_BASE_P, except that that expression may examine the context in which the register appears in the memory reference. outer_code is the code of the immediately enclosing expression (MEM if at the top level of the address, ADDRESS for something that occurs in an address_operand). index_code is the code of the corresponding index expression if outer_code is PLUS; SCRATCH otherwise. The mode may be VOIDmode for addresses that appear outside a MEM, i.e., as an address_operand.

REGNO_OK_FOR_INDEX_P (num) [Macro]
A C expression which is nonzero if register number num is suitable for use as an index register in operand addresses. It may be either a suitable hard register or a pseudo register that has been allocated such a hard register.

The difference between an index register and a base register is that the index register may be scaled. If an address involves the sum of two registers, neither one of them scaled, then either one may be labeled the “base” and the other the “index”; but whichever labeling is used must fit the machine’s constraints of which registers may
serve in each capacity. The compiler will try both labelings, looking for one that is valid, and will reload one or both registers only if neither labeling works.

\[\text{reg\_class\_t TARGET\_PREFERRED\_RENAME\_CLASS (reg\_class\_t \ rclass)}\]  
A target hook that places additional preference on the register class to use when it is necessary to rename a register in class \text{rclass} to another class, or perhaps \text{NO\_REGS}, if no preferred register class is found or hook \text{preferred\_rename\_class} is not implemented. Sometimes returning a more restrictive class makes better code. For example, on ARM, thumb-2 instructions using \text{LO\_REGS} may be smaller than instructions using \text{GENERIC\_REGS}. By returning \text{LO\_REGS} from \text{preferred\_rename\_class}, code size can be reduced.

\[\text{reg\_class\_t TARGET\_PREFERRED\_RELOAD\_CLASS (rtx x, reg\_class\_t \ rclass)}\]  
A target hook that places additional restrictions on the register class to use when it is necessary to copy value \(x\) into a register in class \text{rclass}. The value is a register class; perhaps \text{rclass}, or perhaps another, smaller class.

The default version of this hook always returns value of \text{rclass} argument. Sometimes returning a more restrictive class makes better code. For example, on the 68000, when \(x\) is an integer constant that is in range for a ‘moveq’ instruction, the value of this macro is always \text{DATA\_REGS} as long as \text{rclass} includes the data registers. Requiring a data register guarantees that a ‘moveq’ will be used.

One case where \text{TARGET\_PREFERRED\_RELOAD\_CLASS} must not return \text{rclass} is if \(x\) is a legitimate constant which cannot be loaded into some register class. By returning \text{NO\_REGS} you can force \(x\) into a memory location. For example, rs6000 can load immediate values into general-purpose registers, but does not have an instruction for loading an immediate value into a floating-point register, so \text{TARGET\_PREFERRED\_RELOAD\_CLASS} returns \text{NO\_REGS} when \(x\) is a floating-point constant. If the constant can’t be loaded into any kind of register, code generation will be better if \text{LEGITIMATE\_CONSTANT\_P} makes the constant illegitimate instead of using \text{TARGET\_PREFERRED\_RELOAD\_CLASS}.

If an insn has pseudos in it after register allocation, reload will go through the alternatives and call repeatedly \text{TARGET\_PREFERRED\_RELOAD\_CLASS} to find the best one. Returning \text{NO\_REGS}, in this case, makes reload add a ! in front of the constraint: the x86 back-end uses this feature to discourage usage of 387 registers when math is done in the SSE registers (and vice versa).

\text{PREFERRED\_RELOAD\_CLASS (x, class)}  
A C expression that places additional restrictions on the register class to use when it is necessary to copy value \(x\) into a register in class \text{class}. The value is a register class; perhaps \text{class}, or perhaps another, smaller class. On many machines, the following definition is safe:

\[
\text{#define PREFERRED\_RELOAD\_CLASS(x,CLASS) CLASS}
\]

Sometimes returning a more restrictive class makes better code. For example, on the 68000, when \(x\) is an integer constant that is in range for a ‘moveq’ instruction, the value of this macro is always \text{DATA\_REGS} as long as \text{class} includes the data registers. Requiring a data register guarantees that a ‘moveq’ will be used.
One case where \texttt{PREFERRED\_RELOAD\_CLASS} must not return \texttt{class} is if \(x\) is a legitimate constant which cannot be loaded into some register class. By returning \texttt{NO\_REGS} you can force \(x\) into a memory location. For example, rs6000 can load immediate values into general-purpose registers, but does not have an instruction for loading an immediate value into a floating-point register, so \texttt{PREFERRED\_RELOAD\_CLASS} returns \texttt{NO\_REGS} when \(x\) is a floating-point constant. If the constant can’t be loaded into any kind of register, code generation will be better if \texttt{LEGITIMATE\_CONSTANT\_P} makes the constant illegitimate instead of using \texttt{PREFERRED\_RELOAD\_CLASS}.

If an insn has pseudos in it after register allocation, reload will go through the alternatives and call repeatedly \texttt{PREFERRED\_RELOAD\_CLASS} to find the best one. Returning \texttt{NO\_REGS}, in this case, makes reload add a \! in front of the constraint: the x86 backend uses this feature to discourage usage of 387 registers when math is done in the SSE registers (and vice versa).

\begin{verbatim}
\textbf{PREFERRED\_OUTPUT\_RELOAD\_CLASS} (x, class) \hspace{1cm} \textbf{[Macro]}
Like \texttt{PREFERRED\_RELOAD\_CLASS}, but for output reloads instead of input reloads. If you don’t define this macro, the default is to use \texttt{class}, unchanged.
You can also use \texttt{PREFERRED\_OUTPUT\_RELOAD\_CLASS} to discourage reload from using some alternatives, like \texttt{PREFERRED\_RELOAD\_CLASS}.

\textbf{reg\_class\_t} TARGET\_PREFERRED\_OUTPUT\_RELOAD\_CLASS (rtx x, \hspace{1cm} \textbf{[Target Hook]}
\textbf{reg\_class\_t} rclass)
Like \texttt{TARGET\_PREFERRED\_RELOAD\_CLASS}, but for output reloads instead of input reloads.
The default version of this hook always returns value of \texttt{rclass} argument.
You can also use \texttt{TARGET\_PREFERRED\_OUTPUT\_RELOAD\_CLASS} to discourage reload from using some alternatives, like \texttt{TARGET\_PREFERRED\_RELOAD\_CLASS}.

\textbf{LIMIT\_RELOAD\_CLASS} (mode, class) \hspace{1cm} \textbf{[Macro]}
A C expression that places additional restrictions on the register class to use when it is necessary to be able to hold a value of mode \texttt{mode} in a reload register for which \texttt{class} would ordinarily be used.
Unlike \texttt{PREFERRED\_RELOAD\_CLASS}, this macro should be used when there are certain modes that simply can’t go in certain reload classes.
The value is a register class; perhaps \texttt{class}, or perhaps another, smaller class.
Don’t define this macro unless the target machine has limitations which require the macro to do something nontrivial.

\textbf{reg\_class\_t} TARGET\_SECONDARY\_RELOAD (bool in\_p, rtx x, \hspace{1cm} \textbf{[Target Hook]}
\textbf{reg\_class\_t} reload\_class, enum machine\_mode reload\_mode,
\textbf{secondary\_reload\_info} *sri)
Many machines have some registers that cannot be copied directly to or from memory or even from other types of registers. An example is the ‘MQ’ register, which on most machines, can only be copied to or from general registers, but not memory. Below, we shall be using the term ‘intermediate register’ when a move operation cannot be performed directly, but has to be done by copying the source into the intermediate
register first, and then copying the intermediate register to the destination. An intermediate register always has the same mode as source and destination. Since it holds the actual value being copied, reload might apply optimizations to re-use an intermediate register and eliding the copy from the source when it can determine that the intermediate register still holds the required value.

Another kind of secondary reload is required on some machines which allow copying all registers to and from memory, but require a scratch register for stores to some memory locations (e.g., those with symbolic address on the RT, and those with certain symbolic address on the SPARC when compiling PIC). Scratch registers need not have the same mode as the value being copied, and usually hold a different value than that being copied. Special patterns in the md file are needed to describe how the copy is performed with the help of the scratch register; these patterns also describe the number, register class(es) and mode(s) of the scratch register(s).

In some cases, both an intermediate and a scratch register are required.

For input reloads, this target hook is called with nonzero in_p, and x is an rtx that needs to be copied to a register of class reload_class in reload_mode. For output reloads, this target hook is called with zero in_p, and a register of class reload_class needs to be copied to rtx x in reload_mode.

If copying a register of reload_class from/to x requires an intermediate register, the hook secondary_reload should return the register class required for this intermediate register. If no intermediate register is required, it should return NO_REGS. If more than one intermediate register is required, describe the one that is closest in the copy chain to the reload register.

If scratch registers are needed, you also have to describe how to perform the copy from/to the reload register to/from this closest intermediate register. Or if no intermediate register is required, but still a scratch register is needed, describe the copy from/to the reload register to/from the reload operand x.

You do this by setting sri->icode to the instruction code of a pattern in the md file which performs the move. Operands 0 and 1 are the output and input of this copy, respectively. Operands from operand 2 onward are for scratch operands. These scratch operands must have a mode, and a single-register-class output constraint.

When an intermediate register is used, the secondary_reload hook will be called again to determine how to copy the intermediate register to/from the reload operand x, so your hook must also have code to handle the register class of the intermediate operand.

x might be a pseudo-register or a subreg of a pseudo-register, which could either be in a hard register or in memory. Use true_regnum to find out; it will return -1 if the pseudo is in memory and the hard register number if it is in a register.

Scratch operands in memory (constraint "="/"&") are currently not supported. For the time being, you will have to continue to use SECONDARY_MEMORY_NEEDED for that purpose.

copy_cost also uses this target hook to find out how values are copied. If you want it to include some extra cost for the need to allocate (a) scratch register(s), set sri->extra_cost to the additional cost. Or if two dependent moves are supposed to have a
lower cost than the sum of the individual moves due to expected fortuitous scheduling
and/or special forwarding logic, you can set sri->extra_cost to a negative amount.

```
[Macro]
SECONDARY_RELOAD_CLASS (class, mode, x)
SECONDARY_INPUT_RELOAD_CLASS (class, mode, x)
SECONDARY_OUTPUT_RELOAD_CLASS (class, mode, x)
```

These macros are obsolete, new ports should use the target hook `TARGET_SECONDARY_RELOAD` instead.

These are obsolete macros, replaced by the `TARGET_SECONDARY_RELOAD` target hook. Older ports still define these macros to indicate to the reload phase that it may need to allocate at least one register for a reload in addition to the register to contain the data. Specifically, if copying `x` to a register `class` in `mode` requires an intermediate register, you were supposed to define `SECONDARY_INPUT_RELOAD_CLASS` to return the largest register class all of whose registers can be used as intermediate registers or scratch registers.

If copying a register `class` in `mode` to `x` requires an intermediate or scratch register, `SECONDARY_OUTPUT_RELOAD_CLASS` was supposed to be defined to return the largest register class required. If the requirements for input and output reloads were the same, the macro `SECONDARY_RELOAD_CLASS` should have been used instead of defining both macros identically.

The values returned by these macros are often `GENERAL_REGS`. Return `NO_REGS` if no spare register is needed; i.e., if `x` can be directly copied to or from a register of `class` in `mode` without requiring a scratch register. Do not define this macro if it would always return `NO_REGS`.

If a scratch register is required (either with or without an intermediate register), you were supposed to define patterns for `reload_inm` or `reload_outm`, as required (see Section 16.9 [Standard Names], page 332. These patterns, which were normally implemented with a `define_expand`, should be similar to the `movm` patterns, except that operand 2 is the scratch register.

These patterns need constraints for the reload register and scratch register that contain a single register class. If the original reload register (whose class is `class`) can meet the constraint given in the pattern, the value returned by these macros is used for the class of the scratch register. Otherwise, two additional reload registers are required. Their classes are obtained from the constraints in the insn pattern.

`x` might be a pseudo-register or a `subreg` of a pseudo-register, which could either be in a hard register or in memory. Use `true_regnum` to find out; it will return −1 if the pseudo is in memory and the hard register number if it is in a register.

These macros should not be used in the case where a particular class of registers can only be copied to memory and not to another class of registers. In that case, secondary reload registers are not needed and would not be helpful. Instead, a stack location must be used to perform the copy and the `movm` pattern should use memory as an intermediate storage. This case often occurs between floating-point and general registers.

```
[Macro]
SECONDARY_MEMORY_NEEDED (class1, class2, m)
```

Certain machines have the property that some registers cannot be copied to some other registers without using memory. Define this macro on those machines to be
a C expression that is nonzero if objects of mode \( m \) in registers of \( \text{class1} \) can only be copied to registers of class \( \text{class2} \) by storing a register of \( \text{class1} \) into memory and loading that memory location into a register of \( \text{class2} \).

Do not define this macro if its value would always be zero.

**SECONDARY_MEMORY_NEEDED_RTX (mode)**

[Macro]

Normally when \( \text{SECONDARY_MEMORY_NEEDED} \) is defined, the compiler allocates a stack slot for a memory location needed for register copies. If this macro is defined, the compiler instead uses the memory location defined by this macro.

Do not define this macro if you do not define \( \text{SECONDARY_MEMORY_NEEDED} \).

**SECONDARY_MEMORY_NEEDED_MODE (mode)**

[Macro]

When the compiler needs a secondary memory location to copy between two registers of mode \( \text{mode} \), it normally allocates sufficient memory to hold a quantity of \( \text{BITS\_PER\_WORD} \) bits and performs the store and load operations in a mode that many bits wide and whose class is the same as that of \( \text{mode} \).

This is right thing to do on most machines because it ensures that all bits of the register are copied and prevents accesses to the registers in a narrower mode, which some machines prohibit for floating-point registers.

However, this default behavior is not correct on some machines, such as the DEC Alpha, that store short integers in floating-point registers differently than in integer registers. On those machines, the default widening will not work correctly and you must define this macro to suppress that widening in some cases. See the file ‘alpha.h’ for details.

Do not define this macro if you do not define \( \text{SECONDARY_MEMORY_NEEDED} \) or if widening \( \text{mode} \) to a mode that is \( \text{BITS\_PER\_WORD} \) bits wide is correct for your machine.

**bool TARGET_CLASSLIKELY_SPILLED_P (reg_class_t rclass)**

[Target Hook]

A target hook which returns \( \text{true} \) if pseudos that have been assigned to registers of class \( \text{rclass} \) would likely be spilled because registers of \( \text{rclass} \) are needed for spill registers.

The default version of this target hook returns \( \text{true} \) if \( \text{rclass} \) has exactly one register and \( \text{false} \) otherwise. On most machines, this default should be used. Only use this target hook to some other expression if pseudos allocated by ‘local-alloc.c’ end up in memory because their hard registers were needed for spill registers. If this target hook returns \( \text{false} \) for those classes, those pseudos will only be allocated by ‘global.c’, which knows how to reallocate the pseudo to another register. If there would not be another register available for reallocation, you should not change the implementation of this target hook since the only effect of such implementation would be to slow down register allocation.

**CLASS_MAX_NREGS (class, mode)**

[Macro]

A C expression for the maximum number of consecutive registers of class \( \text{class} \) needed to hold a value of mode \( \text{mode} \).

This is closely related to the macro \( \text{HARD\_REGNO\_NREGS} \). In fact, the value of the macro \( \text{CLASS_MAX_NREGS (class, mode)} \) should be the maximum value of \( \text{HARD\_REGNO\_NREGS (regno, mode)} \) for all \( \text{regno} \) values in the class \( \text{class} \).
This macro helps control the handling of multiple-word values in the reload pass.

**CANNOT_CHANGE_MODE_CLASS (from, to, class)**  
If defined, a C expression that returns nonzero for a class for which a change from mode *from* to mode *to* is invalid.

For the example, loading 32-bit integer or floating-point objects into floating-point registers on the Alpha extends them to 64 bits. Therefore loading a 64-bit object and then storing it as a 32-bit object does not store the low-order 32 bits, as would be the case for a normal register. Therefore, `alpha.h` defines `CANNOT_CHANGE_MODE_CLASS` as below:

```c
#define CANNOT_CHANGE_MODE_CLASS(FROM, TO, CLASS) \
(GET_MODE_SIZE (FROM) != GET_MODE_SIZE (TO) \n    ? reg_classes_intersect_p (FLOAT_REGS, (CLASS)) : 0)
```

**const reg_class_t * TARGET_IRA_COVER_CLASSES (void)**  
Return an array of cover classes for the Integrated Register Allocator (IRA). Cover classes are a set of non-intersecting register classes covering all hard registers used for register allocation purposes. If a move between two registers in the same cover class is possible, it should be cheaper than a load or store of the registers. The array is terminated by a `LIM_REG_CLASSES` element.

The order of cover classes in the array is important. If two classes have the same cost of usage for a pseudo, the class occurred first in the array is chosen for the pseudo.

This hook is called once at compiler startup, after the command-line options have been processed. It is then re-examined by every call to `target_reinit`.

The default implementation returns `IRA_COVER_CLASSES`, if defined, otherwise there is no default implementation. You must define either this macro or `IRA_COVER_CLASSES` in order to use the integrated register allocator with Chaitin-Briggs coloring. If the macro is not defined, the only available coloring algorithm is Chow’s priority coloring.

This hook must not be modified from NULL to non-NULL or vice versa by command-line option processing.

**IRA_COVER_CLASSES**  
See the documentation for `TARGET_IRA_COVER_CLASSES`.

### 17.9 Obsolete Macros for Defining Constraints

Machine-specific constraints can be defined with these macros instead of the machine description constructs described in Section 16.8.7 [Define Constraints], page 329. This mechanism is obsolete. New ports should not use it; old ports should convert to the new mechanism.

**CONSTRAINT_LEN (char, str)**  
For the constraint at the start of *str*, which starts with the letter *c*, return the length. This allows you to have register class / constant / extra constraints that are longer than a single letter; you don’t need to define this macro if you can do with single-letter constraints only. The definition of this macro should use `DEFAULT_CONSTRAINT_LEN` for all the characters that you don’t want to handle specially. There are some sanity checks in genoutput.c that check the constraint
lengths for the md file, so you can also use this macro to help you while you are transitioning from a byzantine single-letter-constraint scheme: when you return a negative length for a constraint you want to re-use, genoutput will complain about every instance where it is used in the md file.

**REG_CLASS_FROM_LETTER (char)**

A C expression which defines the machine-dependent operand constraint letters for register classes. If char is such a letter, the value should be the register class corresponding to it. Otherwise, the value should be NO_REGS. The register letter ‘r’, corresponding to class GENERAL_REGS, will not be passed to this macro; you do not need to handle it.

**REG_CLASS_FROM_CONSTRAINT (char, str)**

Like REG_CLASS_FROM_LETTER, but you also get the constraint string passed in str, so that you can use suffixes to distinguish between different variants.

**CONST_OK_FOR_LETTER_P (value, c)**

A C expression that defines the machine-dependent operand constraint letters (‘I’, ‘J’, ‘K’, . . . ‘P’) that specify particular ranges of integer values. If c is one of those letters, the expression should check that value, an integer, is in the appropriate range and return 1 if so, 0 otherwise. If c is not one of those letters, the value should be 0 regardless of value.

**CONST_OK_FOR_CONSTRAINT_P (value, c, str)**

Like CONST_OK_FOR_LETTER_P, but you also get the constraint string passed in str, so that you can use suffixes to distinguish between different variants.

**CONST_DOUBLE_OK_FOR_LETTER_P (value, c)**

A C expression that defines the machine-dependent operand constraint letters that specify particular ranges of const_double values (‘G’ or ‘H’).

If c is one of those letters, the expression should check that value, an RTX of code const_double, is in the appropriate range and return 1 if so, 0 otherwise. If c is not one of those letters, the value should be 0 regardless of value.

const_double is used for all floating-point constants and for DImode fixed-point constants. A given letter can accept either or both kinds of values. It can use GET_MODE to distinguish between these kinds.

**CONST_DOUBLE_OK_FOR_CONSTRAINT_P (value, c, str)**

Like CONST_DOUBLE_OK_FOR_LETTER_P, but you also get the constraint string passed in str, so that you can use suffixes to distinguish between different variants.

**EXTRA_CONSTRAINT (value, c)**

A C expression that defines the optional machine-dependent constraint letters that can be used to segregate specific types of operands, usually memory references, for the target machine. Any letter that is not elsewhere defined and not matched by REG_CLASS_FROM_LETTER / REG_CLASS_FROM_CONSTRAINT may be used. Normally this macro will not be defined.
If it is required for a particular target machine, it should return 1 if value corresponds to the operand type represented by the constraint letter $c$. If $c$ is not defined as an extra constraint, the value returned should be 0 regardless of value.

For example, on the ROMP, load instructions cannot have their output in r0 if the memory reference contains a symbolic address. Constraint letter ‘Q’ is defined as representing a memory address that does not contain a symbolic address. An alternative is specified with a ‘Q’ constraint on the input and ‘r’ on the output. The next alternative specifies ‘m’ on the input and a register class that does not include r0 on the output.

\[
\text{EXTRA\_CONSTRAINT\_STR (value, c, str)}
\]

Like EXTRA\_CONSTRAINT, but you also get the constraint string passed in str, so that you can use suffixes to distinguish between different variants.

\[
\text{EXTRA\_MEMORY\_CONSTRAINT (c, str)}
\]

A C expression that defines the optional machine-dependent constraint letters, amongst those accepted by EXTRA\_CONSTRAINT, that should be treated like memory constraints by the reload pass.

It should return 1 if the operand type represented by the constraint at the start of str, the first letter of which is the letter $c$, comprises a subset of all memory references including all those whose address is simply a base register. This allows the reload pass to reload an operand, if it does not directly correspond to the operand type of $c$, by copying its address into a base register.

For example, on the S/390, some instructions do not accept arbitrary memory references, but only those that do not make use of an index register. The constraint letter ‘Q’ is defined via EXTRA\_CONSTRAINT as representing a memory address of this type. If the letter ‘Q’ is marked as EXTRA\_MEMORY\_CONSTRAINT, a ‘Q’ constraint can handle any memory operand, because the reload pass knows it can be reloaded by copying the memory address into a base register if required. This is analogous to the way an ‘o’ constraint can handle any memory operand.

\[
\text{EXTRA\_ADDRESS\_CONSTRAINT (c, str)}
\]

A C expression that defines the optional machine-dependent constraint letters, amongst those accepted by EXTRA\_CONSTRAINT / EXTRA\_CONSTRAINT\_STR, that should be treated like address constraints by the reload pass.

It should return 1 if the operand type represented by the constraint at the start of str, which starts with the letter $c$, comprises a subset of all memory addresses including all those that consist of just a base register. This allows the reload pass to reload an operand, if it does not directly correspond to the operand type of str, by copying it into a base register.

Any constraint marked as EXTRA\_ADDRESS\_CONSTRAINT can only be used with the address\_operand predicate. It is treated analogously to the ‘p’ constraint.

### 17.10 Stack Layout and Calling Conventions

This describes the stack layout and calling conventions.
17.10.1 Basic Stack Layout

Here is the basic stack layout.

STACK_GROWS_DOWNWARD [Macro]
Define this macro if pushing a word onto the stack moves the stack pointer to a smaller address.

When we say, “define this macro if . . .”, it means that the compiler checks this macro only with #ifdef so the precise definition used does not matter.

STACK_PUSH_CODE [Macro]
This macro defines the operation used when something is pushed on the stack. In RTL, a push operation will be (set (mem (STACK_PUSH_CODE (reg sp))) ...)
The choices are PRE_DEC, POST_DEC, PRE_INC, and POST_INC. Which of these is correct depends on the stack direction and on whether the stack pointer points to the last item on the stack or whether it points to the space for the next item on the stack.
The default is PRE_DEC when STACK_GROWS_DOWNWARD is defined, which is almost always right, and PRE_INC otherwise, which is often wrong.

FRAME_GROWS_DOWNWARD [Macro]
Define this macro to nonzero value if the addresses of local variable slots are at negative offsets from the frame pointer.

ARGS_GROW_DOWNWARD [Macro]
Define this macro if successive arguments to a function occupy decreasing addresses on the stack.

STARTING_FRAME_OFFSET [Macro]
Offset from the frame pointer to the first local variable slot to be allocated.

If FRAME_GROWS_DOWNWARD, find the next slot’s offset by subtracting the first slot’s length from STARTING_FRAME_OFFSET. Otherwise, it is found by adding the length of the first slot to the value STARTING_FRAME_OFFSET.

STACK_ALIGNMENT_NEEDED [Macro]
Define to zero to disable final alignment of the stack during reload. The nonzero default for this macro is suitable for most ports.

On ports where STARTING_FRAME_OFFSET is nonzero or where there is a register save block following the local block that doesn’t require alignment to STACK_BOUNDARY, it may be beneficial to disable stack alignment and do it in the backend.

STACK_POINTER_OFFSET [Macro]
Offset from the stack pointer register to the first location at which outgoing arguments are placed. If not specified, the default value of zero is used. This is the proper value for most machines.

If ARGS_GROW_DOWNWARD, this is the offset to the location above the first location at which outgoing arguments are placed.
FIRST_PARM_OFFSET \( (\text{fundecl}) \)  
Offset from the argument pointer register to the first argument’s address. On some machines it may depend on the data type of the function.  
If \text{ARGS\_GROW\_DOWNWARD}, this is the offset to the location above the first argument’s address.

STACK_DYNAMIC_OFFSET \( (\text{fundecl}) \)  
Offset from the stack pointer register to an item dynamically allocated on the stack, e.g., by \text{alloca}.

The default value for this macro is \text{STACK\_POINTER\_OFFSET} plus the length of the outgoing arguments. The default is correct for most machines. See ‘\text{function.c}’ for details.

INITIAL_FRAME_ADDRESS_RTX  
A C expression whose value is RTL representing the address of the initial stack frame. This address is passed to \text{RETURN_ADDR_RTX} and \text{DYNAMIC\_CHAIN\_ADDRESS}. If you don’t define this macro, a reasonable default value will be used. Define this macro in order to make frame pointer elimination work in the presence of \_\_builtin\_frame\_address \( (\text{count}) \) and \_\_builtin\_return\_address \( (\text{count}) \) for \text{count} not equal to zero.

DYNAMIC\_CHAIN\_ADDRESS \( (\text{frameaddr}) \)  
A C expression whose value is RTL representing the address in a stack frame where the pointer to the caller’s frame is stored. Assume that \text{frameaddr} is an RTL expression for the address of the stack frame itself.

If you don’t define this macro, the default is to return the value of \text{frameaddr}—that is, the stack frame address is also the address of the stack word that points to the previous frame.

SETUP\_FRAME\_ADDRESSES  
If defined, a C expression that produces the machine-specific code to setup the stack so that arbitrary frames can be accessed. For example, on the SPARC, we must flush all of the register windows to the stack before we can access arbitrary stack frames. You will seldom need to define this macro.

\textbf{rtx TARGET\_BUILTIN\_SETJMP\_FRAME\_VALUE (void)}  
This target hook should return an rtx that is used to store the address of the current frame into the built in \text{setjmp} buffer. The default value, \text{virtual\_stack\_vars\_rtx}, is correct for most machines. One reason you may need to define this target hook is if \text{hard\_frame\_pointer\_rtx} is the appropriate value on your machine.

FRAME\_ADDR\_RTX \( (\text{frameaddr}) \)  
A C expression whose value is RTL representing the value of the frame address for the current frame. \text{frameaddr} is the frame pointer of the current frame. This is used for \_\_builtin\_frame\_address. You need only define this macro if the frame address is not the same as the frame pointer. Most machines do not need to define it.
RETURN_ADDR_RTX (\texttt{count}, \texttt{frameaddr}) \hspace{1cm} \textbf{[Macro]}  
A C expression whose value is RTL representing the value of the return address for the frame \texttt{count} steps up from the current frame, after the prologue. \texttt{frameaddr} is the frame pointer of the \texttt{count} frame, or the frame pointer of the \texttt{count} \(-\) 1 frame if \texttt{RETURN_ADDR_IN_PREVIOUS_FRAME} is defined.

The value of the expression must always be the correct address when \texttt{count} is zero, but may be \texttt{NULL_RTX} if there is no way to determine the return address of other frames.

RETURN_ADDR_IN_PREVIOUS_FRAME \hspace{1cm} \textbf{[Macro]}  
Define this if the return address of a particular stack frame is accessed from the frame pointer of the previous stack frame.

INCOMING_RETURN_ADDR_RTX \hspace{1cm} \textbf{[Macro]}  
A C expression whose value is RTL representing the location of the incoming return address at the beginning of any function, before the prologue. This RTL is either a \texttt{REG}, indicating that the return value is saved in ‘\texttt{REG}’, or a \texttt{MEM} representing a location in the stack.

You only need to define this macro if you want to support call frame debugging information like that provided by DWARF 2.

If this RTL is a \texttt{REG}, you should also define \texttt{DWARF_FRAME_RETURN_COLUMN} to \texttt{DWARF_FRAME_REGNUM (REGNO)}.

DWARF_ALT_FRAME_RETURN_COLUMN \hspace{1cm} \textbf{[Macro]}  
A C expression whose value is an integer giving a DWARF 2 column number that may be used as an alternative return column. The column must not correspond to any gcc hard register (that is, it must not be in the range of \texttt{DWARF_FRAME_REGNUM}).

This macro can be useful if \texttt{DWARF_FRAME_RETURN_COLUMN} is set to a general register, but an alternative column needs to be used for signal frames. Some targets have also used different frame return columns over time.

DWARF_ZERO_REG \hspace{1cm} \textbf{[Macro]}  
A C expression whose value is an integer giving a DWARF 2 register number that is considered to always have the value zero. This should only be defined if the target has an architected zero register, and someone decided it was a good idea to use that register number to terminate the stack backtrace. New ports should avoid this.

\textbf{void TARGET\_DWARF\_HANDLE\_FRAME\_UNSPEC (const char *\texttt{label},  \hspace{1cm} \textbf{[Target Hook]}}  
\texttt{rtx \texttt{pattern}, int \texttt{index})}  
This target hook allows the backend to emit frame-related insns that contain UNSPECS or UNSPEC\_VOLATILEs. The DWARF 2 call frame debugging info engine will invoke it on insns of the form

\begin{verbatim}
(set (reg) (unspec [...] UNSPEC\_INDEX))
\end{verbatim}

and

\begin{verbatim}
(set (reg) (unspec\_volatile [...] UNSPECV\_INDEX)).
\end{verbatim}

to let the backend emit the call frame instructions. \texttt{label} is the CFI label attached to the insn, \texttt{pattern} is the pattern of the insn and \texttt{index} is \texttt{UNSPEC\_INDEX} or \texttt{UNSPECV\_INDEX}.
INCOMING_FRAME_SP_OFFSET  
A C expression whose value is an integer giving the offset, in bytes, from the value of the stack pointer register to the top of the stack frame at the beginning of any function, before the prologue. The top of the frame is defined to be the value of the stack pointer in the previous frame, just before the call instruction.
You only need to define this macro if you want to support call frame debugging information like that provided by DWARF 2.

ARG_POINTER_CFA_OFFSET \((\text{fundecl})\)  
A C expression whose value is an integer giving the offset, in bytes, from the argument pointer to the canonical frame address (cfa). The final value should coincide with that calculated by INCOMING_FRAME_SP_OFFSET. Which is unfortunately not usable during virtual register instantiation.
The default value for this macro is FIRST_PARM_OFFSET \((\text{fundecl})\) + crtl-\>args.pretend_args_size, which is correct for most machines; in general, the arguments are found immediately before the stack frame. Note that this is not the case on some targets that save registers into the caller’s frame, such as SPARC and rs6000, and so such targets need to define this macro.
You only need to define this macro if the default is incorrect, and you want to support call frame debugging information like that provided by DWARF 2.

FRAME_POINTER_CFA_OFFSET \((\text{fundecl})\)  
If defined, a C expression whose value is an integer giving the offset in bytes from the frame pointer to the canonical frame address (cfa). The final value should coincide with that calculated by INCOMING_FRAME_SP_OFFSET.
Normally the CFA is calculated as an offset from the argument pointer, via ARG_POINTER_CFA_OFFSET, but if the argument pointer is variable due to the ABI, this may not be possible. If this macro is defined, it implies that the virtual register instantiation should be based on the frame pointer instead of the argument pointer. Only one of FRAME_POINTER_CFA_OFFSET and ARG_POINTER_CFA_OFFSET should be defined.

CFAFRAME_BASE_OFFSET \((\text{fundecl})\)  
If defined, a C expression whose value is an integer giving the offset in bytes from the canonical frame address (cfa) to the frame base used in DWARF 2 debug information.
The default is zero. A different value may reduce the size of debug information on some ports.

17.10.2 Exception Handling Support

EH RETURN DATA REGNO \((N)\)  
A C expression whose value is the Nth register number used for data by exception handlers, or INVALID_REGNUM if fewer than N registers are usable.
The exception handling library routines communicate with the exception handlers via a set of agreed upon registers. Ideally these registers should be call-clobbered; it is possible to use call-saved registers, but may negatively impact code size. The target must support at least 2 data registers, but should define 4 if there are enough free registers.
You must define this macro if you want to support call frame exception handling like that provided by DWARF 2.

**EH_RETURN_STACKADJ_RTX**  
[Macro]  
A C expression whose value is RTL representing a location in which to store a stack adjustment to be applied before function return. This is used to unwind the stack to an exception handler’s call frame. It will be assigned zero on code paths that return normally.

Typically this is a call-clobbered hard register that is otherwise untouched by the epilogue, but could also be a stack slot.

Do not define this macro if the stack pointer is saved and restored by the regular prolog and epilog code in the call frame itself; in this case, the exception handling library routines will update the stack location to be restored in place. Otherwise, you must define this macro if you want to support call frame exception handling like that provided by DWARF 2.

**EH_RETURN_HANDLER_RTX**  
[Macro]  
A C expression whose value is RTL representing a location in which to store the address of an exception handler to which we should return. It will not be assigned on code paths that return normally.

Typically this is the location in the call frame at which the normal return address is stored. For targets that return by popping an address off the stack, this might be a memory address just below the target call frame rather than inside the current call frame. If defined, **EH_RETURN_STACKADJ_RTX** will have already been assigned, so it may be used to calculate the location of the target call frame.

Some targets have more complex requirements than storing to an address calculable during initial code generation. In that case the **eh_return** instruction pattern should be used instead.

If you want to support call frame exception handling, you must define either this macro or the **eh_return** instruction pattern.

**RETURN_ADDR_OFFSET**  
[Macro]  
If defined, an integer-valued C expression for which rtl will be generated to add it to the exception handler address before it is searched in the exception handling tables, and to subtract it again from the address before using it to return to the exception handler.

**ASM_PREFERRED_EH_DATA_FORMAT**  
[Macro]  
This macro chooses the encoding of pointers embedded in the exception handling sections. If at all possible, this should be defined such that the exception handling section will not require dynamic relocations, and so may be read-only.

code is 0 for data, 1 for code labels, 2 for function pointers. global is true if the symbol may be affected by dynamic relocations. The macro should return a combination of the DW_EH_PE_* defines as found in ‘dwarf2.h’.

If this macro is not defined, pointers will not be encoded but represented directly.
ASM_MAYBE_OUTPUT_ENCODED_ADDR_RTX (file, encoding, size, addr, done)  
This macro allows the target to emit whatever special magic is required to represent  
the encoding chosen by ASM_PREFERRED_EH_DATA_FORMAT. Generic code takes care of  
pc-relative and indirect encodings; this must be defined if the target uses text-relative  
or data-relative encodings.  
This is a C statement that branches to done if the format was handled. encoding is  
the format chosen, size is the number of bytes that the format occupies, addr is the  
SYMBOL_REF to be emitted.

MD_UNWIND_SUPPORT 
A string specifying a file to be #include’d in unwind-dw2.c. The file so included  
typically defines MD_FALLBACK_FRAME_STATE_FOR.

MD_FALLBACK_FRAME_STATE_FOR (context, fs)  
This macro allows the target to add CPU and operating system specific code to  
the call-frame unwinder for use when there is no unwind data available. The most  
common reason to implement this macro is to unwind through signal frames.  
This macro is called from uw_frame_state_for in ‘unwind-dw2.c’,  
‘unwind-dw2-xtensa.c’ and ‘unwind-ia64.c’. context is an _Unwind_Context; fs is an _Unwind_FrameState. Examine context->ra for the address of  
the code being executed and context->cfa for the stack pointer value. If the frame  
can be decoded, the register save addresses should be updated in fs and the macro  
should evaluate to _URC_NO_REASON. If the frame cannot be decoded, the macro  
should evaluate to _URC_END_OF_STACK.  
For proper signal handling in Java this macro is accompanied by MAKE_THROW_FRAME,  
defined in ‘libjava/include/*/signal.h’ headers.

MD_HANDLE_UNWABI (context, fs)  
This macro allows the target to add operating system specific code to the call-frame  
unwinder to handle the IA-64 .unwabi unwinding directive, usually used for signal  
or interrupt frames.  
This macro is called from uw_update_context in ‘unwind-ia64.c’. context is an  
_Unwind_Context; fs is an _Unwind_FrameState. Examine fs->unwabi for the abi  
and context in the .unwabi directive. If the .unwabi directive can be handled, the  
register save addresses should be updated in fs.

TARGET_USES_WEAK_UNWIND_INFO 
A C expression that evaluates to true if the target requires unwind info to be given  
comdat linkage. Define it to be 1 if comdat linkage is necessary. The default is 0.

17.10.3 Specifying How Stack Checking is Done
GCC will check that stack references are within the boundaries of the stack, if the option  
‘-fstack-check’ is specified, in one of three ways:  
1. If the value of the STACK_CHECK_BUILTIN macro is nonzero, GCC will assume that  
you have arranged for full stack checking to be done at appropriate places in the  
configuration files. GCC will not do other special processing.
2. If \texttt{STACK\_CHECK\_BUILTIN} is zero and the value of the \texttt{STACK\_CHECK\_STATIC\_BUILTIN} macro is nonzero, GCC will assume that you have arranged for static stack checking (checking of the static stack frame of functions) to be done at appropriate places in the configuration files. GCC will only emit code to do dynamic stack checking (checking on dynamic stack allocations) using the third approach below.

3. If neither of the above are true, GCC will generate code to periodically “probe” the stack pointer using the values of the macros defined below.

If neither \texttt{STACK\_CHECK\_BUILTIN} nor \texttt{STACK\_CHECK\_STATIC\_BUILTIN} is defined, GCC will change its allocation strategy for large objects if the option ‘-fstack-check’ is specified: they will always be allocated dynamically if their size exceeds \texttt{STACK\_CHECK\_MAX\_VAR\_SIZE} bytes.

\texttt{STACK\_CHECK\_BUILTIN} \hspace{1cm} [Macro]
A nonzero value if stack checking is done by the configuration files in a machine-dependent manner. You should define this macro if stack checking is required by the ABI of your machine or if you would like to do stack checking in some more efficient way than the generic approach. The default value of this macro is zero.

\texttt{STACK\_CHECK\_STATIC\_BUILTIN} \hspace{1cm} [Macro]
A nonzero value if static stack checking is done by the configuration files in a machine-dependent manner. You should define this macro if you would like to do static stack checking in some more efficient way than the generic approach. The default value of this macro is zero.

\texttt{STACK\_CHECK\_PROBE\_INTERVAL\_EXP} \hspace{1cm} [Macro]
An integer specifying the interval at which GCC must generate stack probe instructions, defined as 2 raised to this integer. You will normally define this macro so that the interval be no larger than the size of the “guard pages” at the end of a stack area. The default value of 12 (4096-byte interval) is suitable for most systems.

\texttt{STACK\_CHECK\_MOVING\_SP} \hspace{1cm} [Macro]
An integer which is nonzero if GCC should move the stack pointer page by page when doing probes. This can be necessary on systems where the stack pointer contains the bottom address of the memory area accessible to the executing thread at any point in time. In this situation an alternate signal stack is required in order to be able to recover from a stack overflow. The default value of this macro is zero.

\texttt{STACK\_CHECK\_PROTECT} \hspace{1cm} [Macro]
The number of bytes of stack needed to recover from a stack overflow, for languages where such a recovery is supported. The default value of 75 words with the \texttt{setjmp/longjmp}-based exception handling mechanism and 8192 bytes with other exception handling mechanisms should be adequate for most machines.

The following macros are relevant only if neither \texttt{STACK\_CHECK\_BUILTIN} nor \texttt{STACK\_CHECK\_STATIC\_BUILTIN} is defined; you can omit them altogether in the opposite case.
STACK_CHECK_MAX_FRAME_SIZE

The maximum size of a stack frame, in bytes. GCC will generate probe instructions in non-leaf functions to ensure at least this many bytes of stack are available. If a stack frame is larger than this size, stack checking will not be reliable and GCC will issue a warning. The default is chosen so that GCC only generates one instruction on most systems. You should normally not change the default value of this macro.

STACK_CHECK_FIXED_FRAME_SIZE

GCC uses this value to generate the above warning message. It represents the amount of fixed frame used by a function, not including space for any callee-saved registers, temporaries and user variables. You need only specify an upper bound for this amount and will normally use the default of four words.

STACK_CHECK_MAX_VAR_SIZE

The maximum size, in bytes, of an object that GCC will place in the fixed area of the stack frame when the user specifies `-fstack-check`. GCC computed the default from the values of the above macros and you will normally not need to override that default.

17.10.4 Registers That Address the Stack Frame

This discusses registers that address the stack frame.

STACK_POINTER_REGNUM

The register number of the stack pointer register, which must also be a fixed register according to FIXED_REGISTERS. On most machines, the hardware determines which register this is.

FRAME_POINTER_REGNUM

The register number of the frame pointer register, which is used to access automatic variables in the stack frame. On some machines, the hardware determines which register this is. On other machines, you can choose any register you wish for this purpose.

HARD_FRAME_POINTER_REGNUM

On some machines the offset between the frame pointer and starting offset of the automatic variables is not known until after register allocation has been done (for example, because the saved registers are between these two locations). On those machines, define FRAME_POINTER_REGNUM the number of a special, fixed register to be used internally until the offset is known, and define HARD_FRAME_POINTER_REGNUM to be the actual hard register number used for the frame pointer.

You should define this macro only in the very rare circumstances when it is not possible to calculate the offset between the frame pointer and the automatic variables until after register allocation has been completed. When this macro is defined, you must also indicate in your definition of ELIMINABLE_REGS how to eliminate FRAME_POINTER_REGNUM into either HARD_FRAME_POINTER_REGNUM or STACK_POINTER_REGNUM.

Do not define this macro if it would be the same as FRAME_POINTER_REGNUM.
ARG_POINTER_REGNUM

The register number of the arg pointer register, which is used to access the function’s argument list. On some machines, this is the same as the frame pointer register. On some machines, the hardware determines which register this is. On other machines, you can choose any register you wish for this purpose. If this is not the same register as the frame pointer register, then you must mark it as a fixed register according to FIXED_REGS, or arrange to be able to eliminate it (see Section 17.10.5 [Elimination], page 445).

HARD_FRAME_POINTER_IS_FRAME_POINTER

Define this to a preprocessor constant that is nonzero if hard_frame_pointer_rtx and frame_pointer_rtx should be the same. The default definition is ‘(HARD_FRAME_POINTER_REGNUM == FRAME_POINTER_REGNUM)’; you only need to define this macro if that definition is not suitable for use in preprocessor conditionals.

HARD_FRAME_POINTER_IS_ARG_POINTER

Define this to a preprocessor constant that is nonzero if hard_frame_pointer_rtx and arg_pointer_rtx should be the same. The default definition is ‘(HARD_FRAME_POINTER_REGNUM == ARG_POINTER_REGNUM)’; you only need to define this macro if that definition is not suitable for use in preprocessor conditionals.

RETURN_ADDRESS_POINTER_REGNUM

The register number of the return address pointer register, which is used to access the current function’s return address from the stack. On some machines, the return address is not at a fixed offset from the frame pointer or stack pointer or argument pointer. This register can be defined to point to the return address on the stack, and then be converted by ELIMINABLE_REGS into either the frame pointer or stack pointer. Do not define this macro unless there is no other way to get the return address from the stack.

STATIC_CHAIN_REGNUM

STATIC_CHAIN_INCOMING_REGNUM

Register numbers used for passing a function’s static chain pointer. If register windows are used, the register number as seen by the called function is STATIC_CHAIN_INCOMING_REGNUM, while the register number as seen by the calling function is STATIC_CHAIN_REGNUM. If these registers are the same, STATIC_CHAIN_INCOMING_REGNUM need not be defined.

The static chain register need not be a fixed register.

If the static chain is passed in memory, these macros should not be defined; instead, the TARGET_STATIC_CHAIN hook should be used.

rtx TARGET_STATIC_CHAIN (const tree fndecl, bool incoming_p)

This hook replaces the use of STATIC_CHAIN_REGNUM et al for targets that may use different static chain locations for different nested functions. This may be required if the target has function attributes that affect the calling conventions of the function and those calling conventions use different static chain locations.

The default version of this hook uses STATIC_CHAIN_REGNUM et al.
If the static chain is passed in memory, this hook should be used to provide rtx giving \texttt{mem} expressions that denote where they are stored. Often the \texttt{mem} expression as seen by the caller will be at an offset from the stack pointer and the \texttt{mem} expression as seen by the callee will be at an offset from the frame pointer. The variables \texttt{stack\_pointer\_rtx}, \texttt{frame\_pointer\_rtx}, and \texttt{arg\_pointer\_rtx} will have been initialized and should be used to refer to those items.

\textbf{DWARF\_FRAME\_REGISTERS} \hfill [Macro]
This macro specifies the maximum number of hard registers that can be saved in a call frame. This is used to size data structures used in DWARF2 exception handling.

Prior to GCC 3.0, this macro was needed in order to establish a stable exception handling ABI in the face of adding new hard registers for ISA extensions. In GCC 3.0 and later, the EH ABI is insulated from changes in the number of hard registers. Nevertheless, this macro can still be used to reduce the runtime memory requirements of the exception handling routines, which can be substantial if the ISA contains a lot of registers that are not call-saved.

If this macro is not defined, it defaults to \texttt{FIRST\_PSEUDO\_REGISTER}.

\textbf{PRE\_GCC3\_DWARF\_FRAME\_REGISTERS} \hfill [Macro]
This macro is similar to \texttt{DWARF\_FRAME\_REGISTERS}, but is provided for backward compatibility in pre GCC 3.0 compiled code.

If this macro is not defined, it defaults to \texttt{DWARF\_FRAME\_REGISTERS}.

\textbf{DWARF\_REG\_TO\_UNWIND\_COLUMN} (\texttt{regno}) \hfill [Macro]
Define this macro if the target’s representation for dwarf registers is different than the internal representation for unwind column. Given a dwarf register, this macro should return the internal unwind column number to use instead.

See the PowerPC’s SPE target for an example.

\textbf{DWARF\_FRAME\_REGNUM} (\texttt{regno}) \hfill [Macro]
Define this macro if the target’s representation for dwarf registers used in \texttt{.eh\_frame} or \texttt{.debug\_frame} is different from that used in other debug info sections. Given a GCC hard register number, this macro should return the \texttt{.eh\_frame} register number. The default is \texttt{DBX\_REGISTER\_NUMBER (regno)}.

\textbf{DWARF2\_FRAME\_REG\_OUT} (\texttt{regno, for\_eh}) \hfill [Macro]
Define this macro to map register numbers held in the call frame info that GCC has collected using \texttt{DWARF\_FRAME\_REGNUM} to those that should be output in \texttt{.debug\_frame} (\texttt{for\_eh} is zero) and \texttt{.eh\_frame} (\texttt{for\_eh} is nonzero). The default is to return \texttt{regno}.

17.10.5 Eliminating Frame Pointer and Arg Pointer
This is about eliminating the frame pointer and arg pointer.

\textbf{bool TARGET\_FRAME\_POINTER\_REQUIRED (void)} \hfill [Target Hook]
This target hook should return \texttt{true} if a function must have and use a frame pointer. This target hook is called in the reload pass. If its return value is \texttt{true} the function will have a frame pointer.
This target hook can in principle examine the current function and decide according to the facts, but on most machines the constant `false` or the constant `true` suffices. Use `false` when the machine allows code to be generated with no frame pointer, and doing so saves some time or space. Use `true` when there is no possible advantage to avoiding a frame pointer.

In certain cases, the compiler does not know how to produce valid code without a frame pointer. The compiler recognizes those cases and automatically gives the function a frame pointer regardless of what `TARGET_FRAME_POINTER_REQUIRED` returns. You don’t need to worry about them.

In a function that does not require a frame pointer, the frame pointer register can be allocated for ordinary usage, unless you mark it as a fixed register. See `FIXED_REGS` for more information.

Default return value is `false`.

`INITIAL_FRAME_POINTER_OFFSET (depth-var)`  
A C statement to store in the variable `depth-var` the difference between the frame pointer and the stack pointer values immediately after the function prologue. The value would be computed from information such as the result of `get_frame_size ()` and the tables of registers `regs_ever_live` and `call_used_regs`.

If `ELIMINABLE_REGS` is defined, this macro will be not be used and need not be defined. Otherwise, it must be defined even if `TARGET_FRAME_POINTER_REQUIRED` always returns true; in that case, you may set `depth-var` to anything.

`ELIMINABLE_REGS`  
If defined, this macro specifies a table of register pairs used to eliminate unneeded registers that point into the stack frame. If it is not defined, the only elimination attempted by the compiler is to replace references to the frame pointer with references to the stack pointer.

The definition of this macro is a list of structure initializations, each of which specifies an original and replacement register.

On some machines, the position of the argument pointer is not known until the compilation is completed. In such a case, a separate hard register must be used for the argument pointer. This register can be eliminated by replacing it with either the frame pointer or the argument pointer, depending on whether or not the frame pointer has been eliminated.

In this case, you might specify:

```c
#define ELIMINABLE_REGS  
{{ARG_POINTER_REGNUM, STACK_POINTER_REGNUM}, \ 
 {ARG_POINTER_REGNUM, FRAME_POINTER_REGNUM}, \ 
 {FRAME_POINTER_REGNUM, STACK_POINTER_REGNUM}}
```

Note that the elimination of the argument pointer with the stack pointer is specified first since that is the preferred elimination.

`bool TARGET_CAN_ELIMINATE (const int from_reg, const int to_reg)`  
This target hook should returns `true` if the compiler is allowed to try to replace register number `from_reg` with register number `to_reg`. This target hook need only
be defined if `ELIMINABLE_REGS` is defined, and will usually be `true`, since most of the cases preventing register elimination are things that the compiler already knows about.

Default return value is `true`.

```c
INITIAL_ELIMINATION_OFFSET (from-reg, to-reg, offset-var)  [Macro]
This macro is similar to `INITIAL_FRAME_POINTER_OFFSET`. It specifies the initial difference between the specified pair of registers. This macro must be defined if `ELIMINABLE_REGS` is defined.
```

### 17.10.6 Passing Function Arguments on the Stack

The macros in this section control how arguments are passed on the stack. See the following section for other macros that control passing certain arguments in registers.

```c
bool TARGET_PROMOTE_PROTOTYPES (const tree fntype)  [Target Hook]
This target hook returns `true` if an argument declared in a prototype as an integral type smaller than `int` should actually be passed as an `int`. In addition to avoiding errors in certain cases of mismatch, it also makes for better code on certain machines. The default is to not promote prototypes.
```

```c
PUSH_ARGS  [Macro]
A C expression. If nonzero, push insns will be used to pass outgoing arguments. If the target machine does not have a push instruction, set it to zero. That directs GCC to use an alternate strategy: to allocate the entire argument block and then store the arguments into it. When `PUSH_ARGS` is nonzero, `PUSH_ROUNDING` must be defined too.
```

```c
PUSH_ARGS_REVERSED  [Macro]
A C expression. If nonzero, function arguments will be evaluated from last to first, rather than from first to last. If this macro is not defined, it defaults to `PUSH_ARGS` on targets where the stack and args grow in opposite directions, and 0 otherwise.
```

```c
PUSH_ROUNDING (npushed)  [Macro]
A C expression that is the number of bytes actually pushed onto the stack when an instruction attempts to push `npushed` bytes.
On some machines, the definition
```
#define PUSH_ROUNDING(BYTES) (BYTES)
```

will suffice. But on other machines, instructions that appear to push one byte actually push two bytes in an attempt to maintain alignment. Then the definition should be
```
#define PUSH_ROUNDING(BYTES) (((BYTES) + 1) & ~1)
```

If the value of this macro has a type, it should be an unsigned type.

```c
ACCUMULATE_OUTGOING_ARGS  [Macro]
A C expression. If nonzero, the maximum amount of space required for outgoing arguments will be computed and placed into the variable `current_function_outgoing_args_size`. No space will be pushed onto the stack for each call; instead, the function prologue should increase the stack frame size by this amount.
Setting both `PUSH_ARGS` and `ACCUMULATE_OUTGOING_ARGS` is not proper.
REG_PARM_STACK_SPACE (fndecl)  [Macro]
Define this macro if functions should assume that stack space has been allocated for arguments even when their values are passed in registers.

The value of this macro is the size, in bytes, of the area reserved for arguments passed in registers for the function represented by fndecl, which can be zero if GCC is calling a library function. The argument fndecl can be the FUNCTION_DECL, or the type itself of the function.

This space can be allocated by the caller, or be a part of the machine-dependent stack frame: OUTGOING_REG_PARM_STACK_SPACE says which.

OUTGOING_REG_PARM_STACK_SPACE (fntype)  [Macro]
Define this to a nonzero value if it is the responsibility of the caller to allocate the area reserved for arguments passed in registers when calling a function of fntype. fntype may be NULL if the function called is a library function.

If ACCUMULATE_OUTGOING_ARGS is defined, this macro controls whether the space for these arguments counts in the value of current_function_outgoing_args_size.

STACK_PARMS_IN_REG_PARM_AREA  [Macro]
Define this macro if REG_PARM_STACK_SPACE is defined, but the stack parameters don’t skip the area specified by it.

Normally, when a parameter is not passed in registers, it is placed on the stack beyond the REG_PARM_STACK_SPACE area. Defining this macro suppresses this behavior and causes the parameter to be passed on the stack in its natural location.

int TARGET_RETURN_POPS_ARGS (tree fundecl, tree funtype, int size)  [Target Hook]
This target hook returns the number of bytes of its own arguments that a function pops on returning, or 0 if the function pops no arguments and the caller must therefore pop them all after the function returns.

fundecl is a C variable whose value is a tree node that describes the function in question. Normally it is a node of type FUNCTION_DECL that describes the declaration of the function. From this you can obtain the DECL_ATTRIBUTES of the function.

funtype is a C variable whose value is a tree node that describes the function in question. Normally it is a node of type FUNCTION_TYPE that describes the data type of the function. From this it is possible to obtain the data types of the value and arguments (if known).

When a call to a library function is being considered, fundecl will contain an identifier node for the library function. Thus, if you need to distinguish among various library functions, you can do so by their names. Note that “library function” in this context means a function used to perform arithmetic, whose name is known specially in the compiler and was not mentioned in the C code being compiled.

size is the number of bytes of arguments passed on the stack. If a variable number of bytes is passed, it is zero, and argument popping will always be the responsibility of the calling function.

On the VAX, all functions always pop their arguments, so the definition of this macro is size. On the 68000, using the standard calling convention, no functions pop their
arguments, so the value of the macro is always 0 in this case. But an alternative calling
convention is available in which functions that take a fixed number of arguments pop
them but other functions (such as printf) pop nothing (the caller pops all). When
this convention is in use, functype is examined to determine whether a function takes
a fixed number of arguments.

CALL_POPS_ARGS (cum)  [Macro]
A C expression that should indicate the number of bytes a call sequence pops off the
stack. It is added to the value of RETURN_POPS_ARGS when compiling a function call.
cum is the variable in which all arguments to the called function have been accumu-
lated.

On certain architectures, such as the SH5, a call trampoline is used that pops certain
registers off the stack, depending on the arguments that have been passed to the
function. Since this is a property of the call site, not of the called function, RETURN_
POPS_ARGS is not appropriate.

17.10.7 Passing Arguments in Registers
This section describes the macros which let you control how various types of arguments are
passed in registers or how they are arranged in the stack.

FUNCTION_ARG (cum, mode, type, named)  [Macro]
A C expression that controls whether a function argument is passed in a register, and
which register.

The arguments are cum, which summarizes all the previous arguments; mode, the
machine mode of the argument; type, the data type of the argument as a tree node or
0 if that is not known (which happens for C support library functions); and named,
which is 1 for an ordinary argument and 0 for nameless arguments that correspond to
‘...’ in the called function’s prototype. type can be an incomplete type if a syntax
error has previously occurred.

The value of the expression is usually either a reg RTX for the hard register in which
to pass the argument, or zero to pass the argument on the stack.

For machines like the VAX and 68000, where normally all arguments are pushed, zero
suffices as a definition.

The value of the expression can also be a parallel RTX. This is used when an
argument is passed in multiple locations. The mode of the parallel should be the
mode of the entire argument. The parallel holds any number of expr_list pairs;
each one describes where part of the argument is passed. In each expr_list the
first operand must be a reg RTX for the hard register in which to pass this part
of the argument, and the mode of the register RTX indicates how large this part of
the argument is. The second operand of the expr_list is a const int which gives
the offset in bytes into the entire argument of where this part starts. As a special
exception the first expr_list in the parallel RTX may have a first operand of zero.
This indicates that the entire argument is also stored on the stack.

The last time this macro is called, it is called with MODE == VOIDmode, and its result
is passed to the call or call_value pattern as operands 2 and 3 respectively.
The usual way to make the ISO library `stdarg.h` work on a machine where some arguments are usually passed in registers, is to cause nameless arguments to be passed on the stack instead. This is done by making `FUNCTION_ARG` return 0 whenever `named` is 0.

You may use the hook `targetm.calls.must_pass_in_stack` in the definition of this macro to determine if this argument is of a type that must be passed in the stack. If `REG_PARM_STACK_SPACE` is not defined and `FUNCTION_ARG` returns nonzero for such an argument, the compiler will abort. If `REG_PARM_STACK_SPACE` is defined, the argument will be computed in the stack and then loaded into a register.

```c
bool TARGET_MUST_PASS_IN_STACK (enum machine_mode mode, const tree type) {
    This target hook should return true if we should not pass type solely in registers. The file `stdc++.h` defines a definition that is usually appropriate, refer to `stdc++.h` for additional documentation.
}
```

`FUNCTION_INCOMING_ARG (cum, mode, type, named)`

Define this macro if the target machine has “register windows”, so that the register in which a function sees an arguments is not necessarily the same as the one in which the caller passed the argument.

For such machines, `FUNCTION_ARG` computes the register in which the caller passes the value, and `FUNCTION_INCOMING_ARG` should be defined in a similar fashion to tell the function being called where the arguments will arrive.

If `FUNCTION_INCOMING_ARG` is not defined, `FUNCTION_ARG` serves both purposes.

```c
int TARGET_ARG_PARTIAL_BYTES (CUMULATIVE_ARGS *cum, enum machine_mode mode, tree type, bool named) {
    This target hook returns the number of bytes at the beginning of an argument that must be put in registers. The value must be zero for arguments that are passed entirely in registers or that are entirely pushed on the stack.

    On some machines, certain arguments must be passed partially in registers and partially in memory. On these machines, typically the first few words of arguments are passed in registers, and the rest on the stack. If a multi-word argument (a `double` or a structure) crosses that boundary, its first few words must be passed in registers and the rest must be pushed. This macro tells the compiler when this occurs, and how many bytes should go in registers.

    `FUNCTION_ARG` for these arguments should return the first register to be used by the caller for this argument; likewise `FUNCTION_INCOMING_ARG`, for the called function.
}
```

```c
bool TARGET_PASS_BY_REFERENCE (CUMULATIVE_ARGS *cum, enum machine_mode mode, const tree type, bool named) {
    This target hook should return true if an argument at the position indicated by `cum` should be passed by reference. This predicate is queried after target independent reasons for being passed by reference, such as `TREE_ADDRESSABLE (type)`.

    If the hook returns true, a copy of that argument is made in memory and a pointer to the argument is passed instead of the argument itself. The pointer is passed in whatever way is appropriate for passing a pointer to that type.
}
bool TARGET_CALLEE_COPIES (CUMULATIVE_ARGS *cum, enum machine_mode mode, const tree type, bool named)  

The function argument described by the parameters to this hook is known to be passed by reference. The hook should return true if the function argument should be copied by the callee instead of copied by the caller.

For any argument for which the hook returns true, if it can be determined that the argument is not modified, then a copy need not be generated.

The default version of this hook always returns false.

CUMULATIVE_ARGS  

A C type for declaring a variable that is used as the first argument of FUNCTION_ARG and other related values. For some target machines, the type int suffices and can hold the number of bytes of argument so far.

There is no need to record in CUMULATIVE_ARGS anything about the arguments that have been passed on the stack. The compiler has other variables to keep track of that. For target machines on which all arguments are passed on the stack, there is no need to store anything in CUMULATIVE_ARGS; however, the data structure must exist and should not be empty, so use int.

OVERRIDE_ABI_FORMAT (fndecl)  

If defined, this macro is called before generating any code for a function, but after the cfun descriptor for the function has been created. The back end may use this macro to update cfun to reflect an ABI other than that which would normally be used by default. If the compiler is generating code for a compiler-generated function, fndecl may be NULL.

INIT_CUMULATIVE_ARGS (cum, fntype, libname, fndecl, n_named_args)  

A C statement (sans semicolon) for initializing the variable cum for the state at the beginning of the argument list. The variable has type CUMULATIVE_ARGS. The value of fntype is the tree node for the data type of the function which will receive the args, or 0 if the args are to a compiler support library function. For direct calls that are not libcalls, fndecl contain the declaration node of the function. fndecl is also set when INIT_CUMULATIVE_ARGS is used to find arguments for the function being compiled. n_named_args is set to the number of named arguments, including a structure return address if it is passed as a parameter, when making a call. When processing incoming arguments, n_named_args is set to −1.

When processing a call to a compiler support library function, libname identifies which one. It is a symbol_ref rtx which contains the name of the function, as a string. libname is 0 when an ordinary C function call is being processed. Thus, each time this macro is called, either libname or fntype is nonzero, but never both of them at once.

INIT_CUMULATIVE_LIBCALL_ARGS (cum, mode, libname)  

Like INIT_CUMULATIVE_ARGS but only used for outgoing libcalls, it gets a MODE argument instead of fntype, that would be NULL. indirect would always be zero, too. If this macro is not defined, INIT_CUMULATIVE_ARGS (cum, NULL_RTX, libname, 0) is used instead.
**INIT_CUMULATIVE_INCOMING_ARGS (cum, fntype, libname)**  
Like INIT_CUMULATIVE_ARGS but overrides it for the purposes of finding the arguments for the function being compiled. If this macro is undefined, INIT_CUMULATIVE_ARGS is used instead.

The value passed for libname is always 0, since library routines with special calling conventions are never compiled with GCC. The argument libname exists for symmetry with INIT_CUMULATIVE_ARGS.

**FUNCTION_ARG_ADVANCE (cum, mode, type, named)**  
A C statement (sans semicolon) to update the summarizer variable cum to advance past an argument in the argument list. The values mode, type and named describe that argument. Once this is done, the variable cum is suitable for analyzing the following argument with FUNCTION_ARG, etc.

This macro need not do anything if the argument in question was passed on the stack. The compiler knows how to track the amount of stack space used for arguments without any special help.

**FUNCTION_ARG_OFFSET (mode, type)**  
If defined, a C expression that is the number of bytes to add to the offset of the argument passed in memory. This is needed for the SPU, which passes char and short arguments in the preferred slot that is in the middle of the quad word instead of starting at the top.

**FUNCTION_ARG_PADDING (mode, type)**  
If defined, a C expression which determines whether, and in which direction, to pad out an argument with extra space. The value should be of type enum direction: either upward to pad above the argument, downward to pad below, or none to inhibit padding.

The amount of padding is always just enough to reach the next multiple of TARGET_FUNCTION_ARG_BOUNDARY; this macro does not control it.

This macro has a default definition which is right for most systems. For little-endian machines, the default is to pad upward. For big-endian machines, the default is to pad downward for an argument of constant size shorter than an int, and upward otherwise.

**PAD_VARARGS_DOWN**  
If defined, a C expression which determines whether the default implementation of va_arg will attempt to pad down before reading the next argument, if that argument is smaller than its aligned space as controlled by PARM_BOUNDARY. If this macro is not defined, all such arguments are padded down if BYTES_BIG_ENDIAN is true.

**BLOCK_REG_PADDING (mode, type, first)**  
Specify padding for the last element of a block move between registers and memory. first is nonzero if this is the only element. Defining this macro allows better control of register function parameters on big-endian machines, without using PARALLEL rtl. In particular, MUST_PASS_IN_STACK need not test padding and mode of types in registers, as there is no longer a "wrong" part of a register; For example, a three byte aggregate may be passed in the high part of a register if so required.
unsigned int TARGET_FUNCTION_ARG_BOUNDARY (enum machine_mode mode, const tree type) [Target Hook]
This hook returns the alignment boundary, in bits, of an argument with the specified mode and type. The default hook returns PARM_BOUNDARY for all arguments.

FUNCTION_ARG_REGNO_P (regno) [Macro]
A C expression that is nonzero if regno is the number of a hard register in which function arguments are sometimes passed. This does not include implicit arguments such as the static chain and the structure-value address. On many machines, no registers can be used for this purpose since all function arguments are pushed on the stack.

bool TARGET_SPLIT_COMPLEX_ARG (const tree type) [Target Hook]
This hook should return true if parameter of type type are passed as two scalar parameters. By default, GCC will attempt to pack complex arguments into the target's word size. Some ABIs require complex arguments to be split and treated as their individual components. For example, on AIX64, complex floats should be passed in a pair of floating point registers, even though a complex float would fit in one 64-bit floating point register.

The default value of this hook is NULL, which is treated as always false.

tree TARGET_BUILD_BUILTIN_VA_LIST (void) [Target Hook]
This hook returns a type node for va_list for the target. The default version of the hook returns void*.

int TARGET_ENUM_VA_LIST_P (int idx, const char **pname, tree *ptree) [Target Hook]
This target hook is used in function c_common_nodes_and_builtins to iterate through the target specific builtin types for va_list. The variable idx is used as iterator. pname has to be a pointer to a const char * and ptree a pointer to a tree typed variable. The arguments pname and ptree are used to store the result of this macro and are set to the name of the va_list builtin type and its internal type. If the return value of this macro is zero, then there is no more element. Otherwise the IDX should be increased for the next call of this macro to iterate through all types.

tree TARGET_FN_ABI_VA_LIST (tree fndecl) [Target Hook]
This hook returns the va_list type of the calling convention specified by fndecl. The default version of this hook returns va_list_type_node.

tree TARGET_CANONICAL_VA_LIST_TYPE (tree type) [Target Hook]
This hook returns the va_list type of the calling convention specified by the type of type. If type is not a valid va_list type, it returns NULL_TREE.

tree TARGET_GIMPLIFY_VA_ARG_EXPR (tree valist, tree type, gimple_seq *pre_p, gimple_seq *post_p) [Target Hook]
This hook performs target-specific gimplification of VA_ARG_EXPR. The first two parameters correspond to the arguments to va_arg; the latter two are as in gimplify.c:gimplify_expr.
bool TARGET_VALID_POINTER_MODE (enum machine_mode mode)                   [Target Hook]
    Define this to return nonzero if the port can handle pointers with machine mode
    mode. The default version of this hook returns true for both ptr_mode and Pmode.

bool TARGET_REF_MAY_ALIAS_ERRNO (struct ao_ref.s *ref)                     [Target Hook]
    Define this to return nonzero if the memory reference ref may alias with the system C
    library errno location. The default version of this hook assumes the system C library
    errno location is either a declaration of type int or accessed by dereferencing a pointer
to int.

bool TARGET_SCALAR_MODE_SUPPORTED_P (enum machine_mode mode)               [Target Hook]
    Define this to return nonzero if the port is prepared to handle insns involving scalar
    mode mode. For a scalar mode to be considered supported, all the basic arithmetic
    and comparisons must work.
    The default version of this hook returns true for any mode required to handle the
    basic C types (as defined by the port). Included here are the double-word arithmetic
    supported by the code in ‘optabs.c’.

bool TARGET_VECTOR_MODE_SUPPORTED_P (enum machine_mode mode)               [Target Hook]
    Define this to return nonzero if the port is prepared to handle insns involving vector
    mode mode. At the very least, it must have move patterns for this mode.

bool TARGET_SMALL_REGISTER_CLASSES_FOR_MODE_P (enum machine_mode mode)    [Target Hook]
    Define this to return nonzero for machine modes for which the port has small register
    classes. If this target hook returns nonzero for a given mode, the compiler will try to
    minimize the lifetime of registers in mode. The hook may be called with VOIDmode as
    argument. In this case, the hook is expected to return nonzero if it returns nonzero
    for any mode.
    On some machines, it is risky to let hard registers live across arbitrary insns. Typically,
    these machines have instructions that require values to be in specific registers (like
    an accumulator), and reload will fail if the required hard register is used for another
    purpose across such an insn.
    Passes before reload do not know which hard registers will be used in an instruction,
    but the machine modes of the registers set or used in the instruction are already
    known. And for some machines, register classes are small for, say, integer registers but
    not for floating point registers. For example, the AMD x86-64 architecture requires
    specific registers for the legacy x86 integer instructions, but there are many SSE
    registers for floating point operations. On such targets, a good strategy may be to
    return nonzero from this hook for INTEGRAL_MODE_P machine modes but zero for the
    SSE register classes.
    The default version of this hook returns false for any mode. It is always safe to
    redefine this hook to return with a nonzero value. But if you unnecessarily define it,
you will reduce the amount of optimizations that can be performed in some cases. If
you do not define this hook to return a nonzero value when it is required, the compiler
will run out of spill registers and print a fatal error message.
**17.10.8 How Scalar Function Values Are Returned**

This section discusses the macros that control returning scalars as values—values that can fit in registers.

```c
rtx TARGET_FUNCTION_VALUE (const tree ret_type, const tree fn_decl_or_type, bool outgoing)
```

Define this to return an RTX representing the place where a function returns or receives a value of data type `ret_type`, a tree node representing a data type. `fn_decl_or_type` is a tree node representing `FUNCTION_DECL` or `FUNCTION_TYPE` of a function being called. If `outgoing` is false, the hook should compute the register in which the caller will see the return value. Otherwise, the hook should return an RTX representing the place where a function returns a value.

On many machines, only `TYPE_MODE (ret_type)` is relevant. (Actually, on most machines, scalar values are returned in the same place regardless of mode.) The value of the expression is usually a `reg` RTX for the hard register where the return value is stored. The value can also be a `parallel` RTX, if the return value is in multiple places. See `FUNCTION_ARG` for an explanation of the `parallel` form. Note that the callee will populate every location specified in the `parallel`, but if the first element of the `parallel` contains the whole return value, callers will use that element as the canonical location and ignore the others. The m68k port uses this type of `parallel` to return pointers in both ‘%a0’ (the canonical location) and ‘%d0’.

If `TARGET_PROMOTE_FUNCTION_RETURN` returns true, you must apply the same promotion rules specified in `PROMOTE_MODE` if `valtype` is a scalar type.

If the precise function being called is known, `func` is a tree node (`FUNCTION_DECL`) for it; otherwise, `func` is a null pointer. This makes it possible to use a different value-returning convention for specific functions when all their calls are known.

Some target machines have “register windows” so that the register in which a function returns its value is not the same as the one in which the caller sees the value. For such machines, you should return different RTX depending on `outgoing`.

`TARGET_FUNCTION_VALUE` is not used for return values with aggregate data types, because these are returned in another way. See `TARGET_STRUCT_VALUE_RTX` and related macros, below.

```c
FUNCTION_VALUE (valtype, func)
```

This macro has been deprecated. Use `TARGET_FUNCTION_VALUE` for a new target instead.

```c
LIBCALL_VALUE (mode)
```

A C expression to create an RTX representing the place where a library function returns a value of mode `mode`.

Note that “library function” in this context means a compiler support routine, used to perform arithmetic, whose name is known specially by the compiler and was not mentioned in the C code being compiled.
rtx TARGET_LIBCALL_VALUE (enum machine_mode mode, const rtx fun)  
Define this hook if the back-end needs to know the name of the libcall function in  
or order to determine where the result should be returned.  
The mode of the result is given by mode and the name of the called library function  
is given by fun. The hook should return an RTX representing the place where the  
library function result will be returned.  
If this hook is not defined, then LIBCALL_VALUE will be used.

FUNCTION_VALUE_REGNO_P (regno)  
A C expression that is nonzero if regno is the number of a hard register in which the  
values of called function may come back.  
A register whose use for returning values is limited to serving as the second of a pair  
(for a value of type double, say) need not be recognized by this macro. So for most  
machines, this definition suffices:  

#define FUNCTION_VALUE_REGNO_P(N) ((N) == 0)  
If the machine has register windows, so that the caller and the called function use  
different registers for the return value, this macro should recognize only the caller’s  
register numbers.  
This macro has been deprecated. Use TARGET_FUNCTION_VALUE_REGNO_P for a new  
target instead.

bool TARGET_FUNCTION_VALUE_REGNO_P (const unsigned int regno)  
A target hook that return true if regno is the number of a hard register in which the  
values of called function may come back.  
A register whose use for returning values is limited to serving as the second of a pair  
(for a value of type double, say) need not be recognized by this target hook.  
If the machine has register windows, so that the caller and the called function use  
different registers for the return value, this target hook should recognize only the  
caller’s register numbers.  
If this hook is not defined, then FUNCTION_VALUE_REGNO_P will be used.

APPLY_RESULT_SIZE  
Define this macro if ‘untyped_call’ and ‘untyped_return’ need more space than  
is implied by FUNCTION_VALUE_REGNO_P for saving and restoring an arbitrary return  
value.

bool TARGET_RETURN_IN_MSB (const tree type)  
This hook should return true if values of type type are returned at the most significant  
end of a register (in other words, if they are padded at the least significant end). You  
can assume that type is returned in a register; the caller is required to check this.  
Note that the register provided by TARGET_FUNCTION_VALUE must be able to hold the  
complete return value. For example, if a 1-, 2- or 3-byte structure is returned at the  
most significant end of a 4-byte register, TARGET_FUNCTION_VALUE should provide an  
SImode rtx.
17.10.9 How Large Values Are Returned

When a function value’s mode is BLKmode (and in some other cases), the value is not returned according to TARGET_FUNCTION_VALUE (see Section 17.10.8 [Scalar Return], page 455). Instead, the caller passes the address of a block of memory in which the value should be stored. This address is called the structure value address.

This section describes how to control returning structure values in memory.

**bool TARGET_RETURN_IN_MEMORY (const tree type, const tree fntype)**

This target hook should return a nonzero value to say to return the function value in memory, just as large structures are always returned. Here `type` will be the data type of the value, and `fntype` will be the type of the function doing the returning, or NULL for libcalls.

Note that values of mode BLKmode must be explicitly handled by this function. Also, the option ‘-fpcc-struct-return’ takes effect regardless of this macro. On most systems, it is possible to leave the hook undefined; this causes a default definition to be used, whose value is the constant 1 for BLKmode values, and 0 otherwise.

Do not use this hook to indicate that structures and unions should always be returned in memory. You should instead use DEFAULT_PCC_STRUCT_RETURN to indicate this.

**DEFAULT_PCC_STRUCT_RETURN**

Define this macro to be 1 if all structure and union return values must be in memory. Since this results in slower code, this should be defined only if needed for compatibility with other compilers or with an ABI. If you define this macro to be 0, then the conventions used for structure and union return values are decided by the TARGET_RETURN_IN_MEMORY target hook.

If not defined, this defaults to the value 1.

**rtx TARGET_STRUCT_VALUE_RTX (tree fndecl, int incoming)**

This target hook should return the location of the structure value address (normally a mem or reg), or 0 if the address is passed as an “invisible” first argument. Note that `fndecl` may be NULL, for libcalls. You do not need to define this target hook if the address is always passed as an “invisible” first argument.

On some architectures the place where the structure value address is found by the called function is not the same place that the caller put it. This can be due to register windows, or it could be because the function prologue moves it to a different place. `incoming` is 1 or 2 when the location is needed in the context of the called function, and 0 in the context of the caller.

If `incoming` is nonzero and the address is to be found on the stack, return a mem which refers to the frame pointer. If `incoming` is 2, the result is being used to fetch the structure value address at the beginning of a function. If you need to emit adjusting code, you should do it at this point.

**PCC_STATIC_STRUCT_RETURN**

Define this macro if the usual system convention on the target machine for returning structures and unions is for the called function to return the address of a static variable containing the value.
Do not define this if the usual system convention is for the caller to pass an address to the subroutine. This macro has effect in ‘-fpcc-struct-return’ mode, but it does nothing when you use ‘-freg-struct-return’ mode.

```c
enum machine_mode TARGET_GET_RAW_RESULT_MODE (int regno)  [Target Hook]
This target hook returns the mode to be used when accessing raw return registers in __builtin_return. Define this macro if the value in reg_raw_mode is not correct.
```

```c
enum machine_mode TARGET_GET_RAW_ARG_MODE (int regno)  [Target Hook]
This target hook returns the mode to be used when accessing raw argument registers in __builtin_apply_args. Define this macro if the value in reg_raw_mode is not correct.
```

### 17.10.10 Caller-Saves Register Allocation

If you enable it, GCC can save registers around function calls. This makes it possible to use call-clobbered registers to hold variables that must live across calls.

```c
CALLER_SAVE_PROFITABLE (refs, calls)  [Macro]
A C expression to determine whether it is worthwhile to consider placing a pseudo-register in a call-clobbered hard register and saving and restoring it around each function call. The expression should be 1 when this is worth doing, and 0 otherwise. If you don’t define this macro, a default is used which is good on most machines: 4 * calls < refs.
```

```c
HARD_REGNO_CALLER_SAVE_MODE (regno, nregs)  [Macro]
A C expression specifying which mode is required for saving nregs of a pseudo-register in call-clobbered hard register regno. If regno is unsuitable for caller save, VOIDmode should be returned. For most machines this macro need not be defined since GCC will select the smallest suitable mode.
```

### 17.10.11 Function Entry and Exit

This section describes the macros that output function entry (prologue) and exit (epilogue) code.

```c
void TARGET_ASM_FUNCTION_PROLOGUE (FILE *file,  [Target Hook]
    HOST_WIDE_INT size)
If defined, a function that outputs the assembler code for entry to a function. The prologue is responsible for setting up the stack frame, initializing the frame pointer register, saving registers that must be saved, and allocating size additional bytes of storage for the local variables. size is an integer. file is a stdio stream to which the assembler code should be output.

The label for the beginning of the function need not be output by this macro. That has already been done when the macro is run.

To determine which registers to save, the macro can refer to the array regs_ever_live: element r is nonzero if hard register r is used anywhere within the function. This implies the function prologue should save register r, provided it is not one of the
call-used registers. (TARGET_ASM_FUNCTION_EPILOGUE must likewise use regs-ever-live.)

On machines that have “register windows”, the function entry code does not save on the stack the registers that are in the windows, even if they are supposed to be preserved by function calls; instead it takes appropriate steps to “push” the register stack, if any non-call-used registers are used in the function.

On machines where functions may or may not have frame-pointers, the function entry code must vary accordingly; it must set up the frame pointer if one is wanted, and not otherwise. To determine whether a frame pointer is in wanted, the macro can refer to the variable frame_pointer_needed. The variable’s value will be 1 at run time in a function that needs a frame pointer. See Section 17.10.5 [Elimination], page 445.

The function entry code is responsible for allocating any stack space required for the function. This stack space consists of the regions listed below. In most cases, these regions are allocated in the order listed, with the last listed region closest to the top of the stack (the lowest address if STACK_GROWS_DOWNWARD is defined, and the highest address if it is not defined). You can use a different order for a machine if doing so is more convenient or required for compatibility reasons. Except in cases where required by standard or by a debugger, there is no reason why the stack layout used by GCC need agree with that used by other compilers for a machine.

void TARGET_ASM_FUNCTION_END_PROLOGUE (FILE *file)  
If defined, a function that outputs assembler code at the end of a prologue. This should be used when the function prologue is being emitted as RTL, and you have some extra assembler that needs to be emitted. See [prologue instruction pattern], page 353.

void TARGET_ASM_FUNCTION_BEGIN_EPILOGUE (FILE *file)  
If defined, a function that outputs assembler code at the start of an epilogue. This should be used when the function epilogue is being emitted as RTL, and you have some extra assembler that needs to be emitted. See [epilogue instruction pattern], page 353.

void TARGET_ASM_FUNCTION_EPILOGUE (FILE *file,  
HOST_WIDE_INT size)  
If defined, a function that outputs the assembler code for exit from a function. The epilogue is responsible for restoring the saved registers and stack pointer to their values when the function was called, and returning control to the caller. This macro takes the same arguments as the macro TARGET_ASM_FUNCTION_PROLOGUE, and the registers to restore are determined from regs-ever-live and CALL_USED_REGISTERS in the same way.

On some machines, there is a single instruction that does all the work of returning from the function. On these machines, give that instruction the name ‘return’ and do not define the macro TARGET_ASM_FUNCTION_EPILOGUE at all.

Do not define a pattern named ‘return’ if you want the TARGET_ASM_FUNCTION_EPILOGUE to be used. If you want the target switches to control whether return instructions or epilogues are used, define a ‘return’ pattern with a validity condi-
tion that tests the target switches appropriately. If the ‘return’ pattern’s validity condition is false, epilogues will be used.

On machines where functions may or may not have frame-pointers, the function exit code must vary accordingly. Sometimes the code for these two cases is completely different. To determine whether a frame pointer is wanted, the macro can refer to the variable `frame_pointer_needed`. The variable’s value will be 1 when compiling a function that needs a frame pointer.

Normally, `TARGET_ASM_FUNCTION_PROLOGUE` and `TARGET_ASM_FUNCTION_EPILOGUE` must treat leaf functions specially. The C variable `current_function_is_leaf` is nonzero for such a function. See Section 17.7.4 [Leaf Functions], page 423.

On some machines, some functions pop their arguments on exit while others leave that for the caller to do. For example, the 68020 when given ‘-mrtl’ pops arguments in functions that take a fixed number of arguments.

Your definition of the macro `RETURN_POPS_ARGS` decides which functions pop their own arguments. `TARGET_ASM_FUNCTION_EPILOGUE` needs to know what was decided. The number of bytes of the current function’s arguments that this function should pop is available in `crt1->args.pops_args`. See Section 17.10.8 [Scalar Return], page 455.

- A region of `current_function_pretend_args_size` bytes of uninitialized space just underneath the first argument arriving on the stack. (This may not be at the very start of the allocated stack region if the calling sequence has pushed anything else since pushing the stack arguments. But usually, on such machines, nothing else has been pushed yet, because the function prologue itself does all the pushing.) This region is used on machines where an argument may be passed partly in registers and partly in memory, and, in some cases to support the features in `<stdarg.h>`.

- An area of memory used to save certain registers used by the function. The size of this area, which may also include space for such things as the return address and pointers to previous stack frames, is machine-specific and usually depends on which registers have been used in the function. Machines with register windows often do not require a save area.

- A region of at least `size` bytes, possibly rounded up to an allocation boundary, to contain the local variables of the function. On some machines, this region and the save area may occur in the opposite order, with the save area closer to the top of the stack.

- Optionally, when `ACCUMULATE_OUTGOING_ARGS` is defined, a region of `current_function_outgoing_args_size` bytes to be used for outgoing argument lists of the function. See Section 17.10.6 [Stack Arguments], page 447.

**EXIT_IGNORE_STACK** [Macro]

Define this macro as a C expression that is nonzero if the return instruction or the function epilogue ignores the value of the stack pointer; in other words, if it is safe to delete an instruction to adjust the stack pointer before a return from the function. The default is 0.

Note that this macro’s value is relevant only for functions for which frame pointers are maintained. It is never safe to delete a final stack adjustment in a function that has no frame pointer, and the compiler knows this regardless of `EXIT_IGNORE_STACK`. 
Chapter 17: Target Description Macros and Functions

**EPILOGUE_USES** *(regno)*

Define this macro as a C expression that is nonzero for registers that are used by the epilogue or the 'return' pattern. The stack and frame pointer registers are already assumed to be used as needed.

**EH_USES** *(regno)*

Define this macro as a C expression that is nonzero for registers that are used by the exception handling mechanism, and so should be considered live on entry to an exception edge.

**DELAY_SLOTS_FOR_EPILOGUE**

Define this macro if the function epilogue contains delay slots to which instructions from the rest of the function can be “moved”. The definition should be a C expression whose value is an integer representing the number of delay slots there.

**ELIGIBLE_FOR_EPILOGUE_DELAY** *(insn, n)*

A C expression that returns 1 if insn can be placed in delay slot number n of the epilogue.

The argument n is an integer which identifies the delay slot now being considered (since different slots may have different rules of eligibility). It is never negative and is always less than the number of epilogue delay slots (what **DELAY_SLOTS_FOR_EPILOGUE** returns). If you reject a particular insn for a given delay slot, in principle, it may be reconsidered for a subsequent delay slot. Also, other insns may (at least in principle) be considered for the so far unfilled delay slot.

The insns accepted to fill the epilogue delay slots are put in an RTL list made with **insn_list** objects, stored in the variable **current_function_epilogue_delay_list**. The insn for the first delay slot comes first in the list. Your definition of the macro **TARGET_ASM_FUNCTION_EPILOGUE** should fill the delay slots by outputting the insns in this list, usually by calling **final_scan_insn**.

You need not define this macro if you did not define **DELAY_SLOTS_FOR_EPILOGUE**.

**TARGET_ASM_OUTPUT_MI_THUNK** *(FILE *file, tree thunk_fndecl, HOST_WIDE_INT delta, HOST_WIDE_INT vcall_offset, tree function)*

A function that outputs the assembler code for a thunk function, used to implement C++ virtual function calls with multiple inheritance. The thunk acts as a wrapper around a virtual function, adjusting the implicit object parameter before handing control off to the real function.

First, emit code to add the integer delta to the location that contains the incoming first argument. Assume that this argument contains a pointer, and is the one used to pass the this pointer in C++. This is the incoming argument before the function prologue, e.g. ‘%o0’ on a sparc. The addition must preserve the values of all other incoming arguments.

Then, if **vcall_offset** is nonzero, an additional adjustment should be made after adding delta. In particular, if p is the adjusted pointer, the following adjustment should be made:
p += (*((ptrdiff_t **)p))[vcall_offset/sizeof(ptrdiff_t)]

After the additions, emit code to jump to `function`, which is a `FUNCTION_DECL`. This is a direct pure jump, not a call, and does not touch the return address. Hence returning from `FUNCTION` will return to whoever called the current ‘thunk’.

The effect must be as if `function` had been called directly with the adjusted first argument. This macro is responsible for emitting all of the code for a thunk function; `TARGET_ASM_FUNCTION_PROLOGUE` and `TARGET_ASM_FUNCTION_EPILOGUE` are not invoked.

The `thunk_fndecl` is redundant. (`delta` and `function` have already been extracted from it.) It might possibly be useful on some targets, but probably not.

If you do not define this macro, the target-independent code in the C++ front end will generate a less efficient heavyweight thunk that calls `function` instead of jumping to it. The generic approach does not support varargs.

```
bool TARGET_ASM_CAN_OUTPUT_MI_THUNK (const_tree thunk_fndecl, HOST_WIDE_INT delta, HOST_WIDE_INT vcall_offset, const_tree function) [Target Hook]
A function that returns true if `TARGET_ASM_OUTPUT_MI_THUNK` would be able to output the assembler code for the thunk function specified by the arguments it is passed, and false otherwise. In the latter case, the generic approach will be used by the C++ front end, with the limitations previously exposed.
```

### 17.10.12 Generating Code for Profiling

These macros will help you generate code for profiling.

```
FUNCTION_PROFILER (file, labelno) [Macro]
A C statement or compound statement to output to `file` some assembler code to call the profiling subroutine `mcount`.

The details of how `mcount` expects to be called are determined by your operating system environment, not by GCC. To figure them out, compile a small program for profiling using the system’s installed C compiler and look at the assembler code that results.

Older implementations of `mcount` expect the address of a counter variable to be loaded into some register. The name of this variable is ‘LP’ followed by the number `labelno`, so you would generate the name using ‘LP%d’ in a `fprintf`.

PROFILE_HOOK [Macro]
A C statement or compound statement to output to `file` some assembly code to call the profiling subroutine `mcount` even the target does not support profiling.

NO_PROFILE_COUNTERS [Macro]
Define this macro to be an expression with a nonzero value if the `mcount` subroutine on your system does not need a counter variable allocated for each function. This is true for almost all modern implementations. If you define this macro, you must not use the `labelno` argument to `FUNCTION_PROFILER`.
```
 PROFILE_BEFORE_PROLOGUE

Define this macro if the code for function profiling should come before the function prologue. Normally, the profiling code comes after.

17.10.13 Permitting tail calls

bool TARGET_FUNCTION_OK_FOR_SIBCALL (tree decl, tree exp) [Target Hook]

True if it is ok to do sibling call optimization for the specified call expression exp. decl will be the called function, or NULL if this is an indirect call.

It is not uncommon for limitations of calling conventions to prevent tail calls to functions outside the current unit of translation, or during PIC compilation. The hook is used to enforce these restrictions, as the sibcall md pattern can not fail, or fall over to a “normal” call. The criteria for successful sibling call optimization may vary greatly between different architectures.

void TARGET_EXTRA_LIVE_ON_ENTRY (bitmap regs) [Target Hook]

Add any hard registers to regs that are live on entry to the function. This hook only needs to be defined to provide registers that cannot be found by examination of FUNCTION_ARG_REGNO_P, the callee saved registers, STATIC_CHAIN_INCOMING_REGNUM, STATIC_CHAIN_REGNUM, TARGET_STRUCT_VALUE_RTX, FRAME_POINTER_REGNUM, EHUSES, FRAME_POINTER_REGNUM, ARG_POINTER_REGNUM, and the PIC_OFFSET_TABLE_REGNUM.

17.10.14 Stack smashing protection

tree TARGET_STACK_PROTECT_GUARD (void) [Target Hook]

This hook returns a DECL node for the external variable to use for the stack protection guard. This variable is initialized by the runtime to some random value and is used to initialize the guard value that is placed at the top of the local stack frame. The type of this variable must be ptr_type_node.

The default version of this hook creates a variable called ‘__stack_chk_guard’, which is normally defined in ‘libgcc2.c’.

tree TARGET_STACK_PROTECT_FAIL (void) [Target Hook]

This hook returns a tree expression that alerts the runtime that the stack protect guard variable has been modified. This expression should involve a call to a noreturn function.

The default version of this hook invokes a function called ‘__stack_chk_fail’, taking no arguments. This function is normally defined in ‘libgcc2.c’.

bool TARGET_SUPPORTS_SPLIT_STACK (bool report, struct gcc_options *opts) [Target Hook]

Whether this target supports splitting the stack when the options described in opts have been passed. This is called after options have been parsed, so the target may reject splitting the stack in some configurations. The default version of this hook returns false. If report is true, this function may issue a warning or error; if report is false, it must simply return a value.
17.11 Implementing the Varargs Macros

GCC comes with an implementation of `<varargs.h>` and `<stdarg.h>` that work without change on machines that pass arguments on the stack. Other machines require their own implementations of varargs, and the two machine independent header files must have conditionals to include it.

ISO `<stdarg.h>` differs from traditional `<varargs.h>` mainly in the calling convention for `va_start`. The traditional implementation takes just one argument, which is the variable in which to store the argument pointer. The ISO implementation of `va_start` takes an additional second argument. The user is supposed to write the last named argument of the function here.

However, `va_start` should not use this argument. The way to find the end of the named arguments is with the built-in functions described below.

```c
__builtin_savereg ()
```

[Macro]

Use this built-in function to save the argument registers in memory so that the varargs mechanism can access them. Both ISO and traditional versions of `va_start` must use `__builtin_savereg`, unless you use `TARGET_SETUP_INCOMING_VARARGS` (see below) instead.

On some machines, `__builtin_savereg` is open-coded under the control of the target hook `TARGET_EXPAND_BUILTIN_SAVEREGRS`. On other machines, it calls a routine written in assembler language, found in `libgcc2.c`.

Code generated for the call to `__builtin_savereg` appears at the beginning of the function, as opposed to where the call to `__builtin_savereg` is written, regardless of what the code is. This is because the registers must be saved before the function starts to use them for its own purposes.

```c
__builtin_next_arg (lastarg)
```

[Macro]

This built-in returns the address of the first anonymous stack argument, as type `void *`. If `ARGS_GROW_DOWNWARD`, it returns the address of the location above the first anonymous stack argument. Use it in `va_start` to initialize the pointer for fetching arguments from the stack. Also use it in `va_start` to verify that the second parameter `lastarg` is the last named argument of the current function.

```c
__builtin_classify_type (object)
```

[Macro]

Since each machine has its own conventions for which data types are passed in which kind of register, your implementation of `va_arg` has to embody these conventions. The easiest way to categorize the specified data type is to use `__builtin_classify_type` together with `sizeof` and `__alignof__`.

`__builtin_classify_type` ignores the value of `object`, considering only its data type. It returns an integer describing what kind of type that is—integer, floating, pointer, structure, and so on.

The file `typeclass.h` defines an enumeration that you can use to interpret the values of `__builtin_classify_type`.

These machine description macros help implement varargs:
rtx TARGET_EXPAND_BUILTIN_SAVEREGS (void) [Target Hook]
If defined, this hook produces the machine-specific code for a call to __builtin_savereg. This code will be moved to the very beginning of the function, before any parameter access are made. The return value of this function should be an RTX that contains the value to use as the return of __builtin_savereg.

void TARGET_SETUP_INCOMING_VARARGS (CUMULATIVE_ARGS *args_so_far, enum machine_mode mode, tree type, int *pretend_args_size, int second_time) [Target Hook]
This target hook offers an alternative to using __builtin_savereg and defining the hook TARGET_EXPAND_BUILTIN_SAVEREGS. Use it to store the anonymous register arguments into the stack so that all the arguments appear to have been passed consecutively on the stack. Once this is done, you can use the standard implementation of varargs that works for machines that pass all their arguments on the stack.

The argument args_so_far points to the CUMULATIVE_ARGS data structure, containing the values that are obtained after processing the named arguments. The arguments mode and type describe the last named argument—its machine mode and its data type as a tree node.

The target hook should do two things: first, push onto the stack all the argument registers not used for the named arguments, and second, store the size of the data thus pushed into the int-valued variable pointed to by pretend_args_size. The value that you store here will serve as additional offset for setting up the stack frame.

Because you must generate code to push the anonymous arguments at compile time without knowing their data types, TARGET_SETUP_INCOMING_VARARGS is only useful on machines that have just a single category of argument register and use it uniformly for all data types.

If the argument second_time is nonzero, it means that the arguments of the function are being analyzed for the second time. This happens for an inline function, which is not actually compiled until the end of the source file. The hook TARGET_SETUP_INCOMING_VARARGS should not generate any instructions in this case.

bool TARGET_STRICT_ARGUMENT_NAMING (CUMULATIVE_ARGS *ca) [Target Hook]
Define this hook to return true if the location where a function argument is passed depends on whether or not it is a named argument.

This hook controls how the named argument to FUNCTION_ARG is set for varargs andstdarg functions. If this hook returns true, the named argument is always true for named arguments, and false for unnamed arguments. If it returns false, but TARGET_PRETEND_OUTGOING_VARARGS_NAMED returns true, then all arguments are treated as named. Otherwise, all named arguments except the last are treated as named.

You need not define this hook if it always returns false.

bool TARGET_PRETEND_OUTGOING_VARARGS_NAMED (CUMULATIVE_ARGS *ca) [Target Hook]
If you need to conditionally change ABIs so that one works with TARGET_SETUP_INCOMING_VARARGS, but the other works like neither TARGET_SETUP_INCOMING_VARARGS nor TARGET_STRICT_ARGUMENT_NAMING was defined, then define this hook.
to return \texttt{true} if \texttt{TARGET\_SETUP\_INCOMING\_VARARGS} is used, \texttt{false} otherwise. Otherwise, you should not define this hook.

17.12 Trampolines for Nested Functions

A \textit{trampoline} is a small piece of code that is created at run time when the address of a nested function is taken. It normally resides on the stack, in the stack frame of the containing function. These macros tell GCC how to generate code to allocate and initialize a trampoline.

The instructions in the trampoline must do two things: load a constant address into the static chain register, and jump to the real address of the nested function. On CISC machines such as the m68k, this requires two instructions, a move immediate and a jump. Then the two addresses exist in the trampoline as word-long immediate operands. On RISC machines, it is often necessary to load each address into a register in two parts. Then pieces of each address form separate immediate operands.

The code generated to initialize the trampoline must store the variable parts—the static chain value and the function address—into the immediate operands of the instructions. On a CISC machine, this is simply a matter of copying each address to a memory reference at the proper offset from the start of the trampoline. On a RISC machine, it may be necessary to take out pieces of the address and store them separately.

\begin{verbatim}
void TARGET_ASM_TRAMPOLINE_TEMPLATE (FILE *f) { [Target Hook]
This hook is called by \texttt{assemble_trampoline_template} to output, on the stream \texttt{f}, assembler code for a block of data that contains the constant parts of a trampoline. This code should not include a label—the label is taken care of automatically.

If you do not define this hook, it means no template is needed for the target. Do not define this hook on systems where the block move code to copy the trampoline into place would be larger than the code to generate it on the spot.

TRAMPOLINE\_SECTION [Macro]
Return the section into which the trampoline template is to be placed (see Section 17.19 \[Sections\], page 491). The default value is \texttt{readonly\_data\_section}.

TRAMPOLINE\_SIZE [Macro]
A C expression for the size in bytes of the trampoline, as an integer.

TRAMPOLINE\_ALIGNMENT [Macro]
Alignment required for trampolines, in bits.

If you don’t define this macro, the value of \texttt{FUNCTION\_ALIGNMENT} is used for aligning trampolines.

void TARGET_TRAMPOLINE\_INIT (rtx m_tramp, tree fndecl, rtx static\_chain) { [Target Hook]
This hook is called to initialize a trampoline. \texttt{m_tramp} is an RTX for the memory block for the trampoline; \texttt{fndecl} is the \texttt{FUNCTION\_DECL} for the nested function; \texttt{static\_chain} is an RTX for the static chain value that should be passed to the function when it is called.
\end{verbatim}
If the target defines `TARGET_ASM_TRAMPOLINE_TEMPLATE`, then the first thing this hook should do is emit a block move into `m_tramp` from the memory block returned by `assemble_trampoline_template`. Note that the block move need only cover the constant parts of the trampoline. If the target isolates the variable parts of the trampoline to the end, not all `TRAMPOLINE_SIZE` bytes need be copied.

If the target requires any other actions, such as flushing caches or enabling stack execution, these actions should be performed after initializing the trampoline proper.

```c
rtx TARGET_TRAMPOLINE_ADJUST_ADDRESS (rtx addr)  [Target Hook]
```

This hook should perform any machine-specific adjustment in the address of the trampoline. Its argument contains the address of the memory block that was passed to `TARGET_TRAMPOLINE_INIT`. In case the address to be used for a function call should be different from the address at which the template was stored, the different address should be returned; otherwise `addr` should be returned unchanged. If this hook is not defined, `addr` will be used for function calls.

Implementing trampolines is difficult on many machines because they have separate instruction and data caches. Writing into a stack location fails to clear the memory in the instruction cache, so when the program jumps to that location, it executes the old contents.

Here are two possible solutions. One is to clear the relevant parts of the instruction cache whenever a trampoline is set up. The other is to make all trampolines identical, by having them jump to a standard subroutine. The former technique makes trampoline execution faster; the latter makes initialization faster.

To clear the instruction cache when a trampoline is initialized, define the following macro.

```c
CLEAR_INSN_CACHE (beg, end)  [Macro]
```

If defined, expands to a C expression clearing the instruction cache in the specified interval. The definition of this macro would typically be a series of `asm` statements. Both `beg` and `end` are both pointer expressions.

The operating system may also require the stack to be made executable before calling the trampoline. To implement this requirement, define the following macro.

```c
ENABLE_EXECUTE_STACK  [Macro]
```

Define this macro if certain operations must be performed before executing code located on the stack. The macro should expand to a series of C file-scope constructs (e.g. functions) and provide a unique entry point named `__enable_execute_stack`. The target is responsible for emitting calls to the entry point in the code, for example from the `TARGET_TRAMPOLINE_INIT` hook.

To use a standard subroutine, define the following macro. In addition, you must make sure that the instructions in a trampoline fill an entire cache line with identical instructions, or else ensure that the beginning of the trampoline code is always aligned at the same point in its cache line. Look in ‘m68k.h’ as a guide.

```c
TRANSFER_FROM_TRAMPOLINE  [Macro]
```

Define this macro if trampolines need a special subroutine to do their work. The macro should expand to a series of `asm` statements which will be compiled with GCC. They go in a library function named `__transfer_from_trampoline`.
If you need to avoid executing the ordinary prologue code of a compiled C function when you jump to the subroutine, you can do so by placing a special label of your own in the assembler code. Use one `asm` statement to generate an assembler label, and another to make the label global. Then trampolines can use that label to jump directly to your special assembler code.

### 17.13 Implicit Calls to Library Routines

Here is an explanation of implicit calls to library routines.

**DECLARE_LIBRARY_RENAMES**  
This macro, if defined, should expand to a piece of C code that will get expanded when compiling functions for libgcc.a. It can be used to provide alternate names for GCC’s internal library functions if there are ABI-mandated names that the compiler should provide.

```c
void TARGET_INIT_LIBFUNCS (void)  
This hook should declare additional library routines or rename existing ones, using the functions `set_optab_libfunc` and `init_one_libfunc` defined in ‘optabs.c’. `init_optabs` calls this macro after initializing all the normal library routines.

The default is to do nothing. Most ports don’t need to define this hook.

**FLOAT_LIB_COMPARE_RETURNS_BOOL**  
This macro should return `true` if the library routine that implements the floating point comparison operator `comparison` in mode `mode` will return a boolean, and `false` if it will return a tristate.

GCC’s own floating point libraries return tristates from the comparison operators, so the default returns false always. Most ports don’t need to define this macro.

**TARGET_LIB_INT_CMP_BIASED**  
This macro should evaluate to `true` if the integer comparison functions (like `__cmpdi2`) return 0 to indicate that the first operand is smaller than the second, 1 to indicate that they are equal, and 2 to indicate that the first operand is greater than the second. If this macro evaluates to `false` the comparison functions return −1, 0, and 1 instead of 0, 1, and 2. If the target uses the routines in ‘libgcc.a’, you do not need to define this macro.

**TARGET_EDOM**  
The value of `EDOM` on the target machine, as a C integer constant expression. If you don’t define this macro, GCC does not attempt to deposit the value of `EDOM` into `errno` directly. Look in ‘/usr/include/errno.h’ to find the value of `EDOM` on your system.

If you do not define `TARGET_EDOM`, then compiled code reports domain errors by calling the library function and letting it report the error. If mathematical functions on your system use `matherr` when there is an error, then you should leave `TARGET_EDOM` undefined so that `matherr` is used normally.
GEN_ERRNO_RTX

[Macro]

Define this macro as a C expression to create an rtl expression that refers to the
global “variable” errno. (On certain systems, errno may not actually be a variable.)
If you don’t define this macro, a reasonable default is used.

TARGET_C99_FUNCTIONS

[Macro]

When this macro is nonzero, GCC will implicitly optimize sin calls into sinf and
similarly for other functions defined by C99 standard. The default is zero because a
number of existing systems lack support for these functions in their runtime so this
macro needs to be redefined to one on systems that do support the C99 runtime.

TARGET_HAS_SINCOS

[Macro]

When this macro is nonzero, GCC will implicitly optimize calls to sin and cos with
the same argument to a call to sincos. The default is zero. The target has to provide
the following functions:

```c
void sincos(double x, double *sin, double *cos);
void sincosf(float x, float *sin, float *cos);
void sincosl(long double x, long double *sin, long double *cos);
```

NEXT_OBJC_RUNTIME

[Macro]

Define this macro to generate code for Objective-C message sending using the calling
convention of the NeXT system. This calling convention involves passing the object,
the selector and the method arguments all at once to the method-lookup library
function.

The default calling convention passes just the object and the selector to the lookup
function, which returns a pointer to the method.

17.14 Addressing Modes

This is about addressing modes.

HAVE_PRE_INCREMENT

[Macro]

HAVE_PRE_DECREMENT

[Macro]

HAVE_POST_INCREMENT

[Macro]

HAVE_POST_DECREMENT

[Macro]

A C expression that is nonzero if the machine supports pre-increment, pre-decrement,
post-increment, or post-decrement addressing respectively.

HAVE_PRE_MODIFY_DISP

[Macro]

HAVE_POST_MODIFY_DISP

[Macro]

A C expression that is nonzero if the machine supports pre- or post-address side-effect
generation involving constants other than the size of the memory operand.

HAVE_PRE_MODIFY_REG

[Macro]

HAVE_POST_MODIFY_REG

[Macro]

A C expression that is nonzero if the machine supports pre- or post-address side-effect
generation involving a register displacement.
CONSTANT_ADDRESS_P (x)  [Macro]
A C expression that is 1 if the RTX x is a constant which is a valid address.
On most machines the default definition of (CONSTANT_P (x) && GET_CODE (x) !=
CONST_DOUBLE) is acceptable, but a few machines are more restrictive as to which
constant addresses are supported.

CONSTANT_P (x)  [Macro]
CONSTANT_P, which is defined by target-independent code, accepts integer-values ex-
pressions whose values are not explicitly known, such as symbol_ref, label_ref,
and high expressions and const arithmetic expressions, in addition to const_int
and const_double expressions.

MAX_REGS_PER_ADDRESS  [Macro]
A number, the maximum number of registers that can appear in a valid memory
address. Note that it is up to you to specify a value equal to the maximum number
that TARGET_LEGITIMATE_ADDRESS_P would ever accept.

bool TARGET_LEGITIMATE_ADDRESS_P (enum machine_mode mode,
rtx x, bool strict)  [Target Hook]
A function that returns whether x (an RTX) is a legitimate memory address on the
target machine for a memory operand of mode mode.
Legitimate addresses are defined in two variants: a strict variant and a non-strict one. The strict parameter chooses which variant is desired by the caller.
The strict variant is used in the reload pass. It must be defined so that any pseudo-
register that has not been allocated a hard register is considered a memory reference. This is because in contexts where some kind of register is required, a pseudo-register with no hard register must be rejected. For non-hard registers, the strict variant
should look up the reg_renumber array; it should then proceed using the hard register
number in the array, or treat the pseudo as a memory reference if the array holds -1.
The non-strict variant is used in other passes. It must be defined to accept all pseudo-
registers in every context where some kind of register is required.
Normally, constant addresses which are the sum of a symbol_ref and an integer are
stored inside a const RTX to mark them as constant. Therefore, there is no need to
recognize such sums specifically as legitimate addresses. Normally you would simply
recognize any const as legitimate.

Usually PRINT_OPERAND_ADDRESS is not prepared to handle constant sums that are
not marked with const. It assumes that a naked plus indicates indexing. If so, then
you must reject such naked constant sums as illegitimate addresses, so that none of
them will be given to PRINT_OPERAND_ADDRESS.

On some machines, whether a symbolic address is legitimate depends on the section
that the address refers to. On these machines, define the target hook TARGET_ENCODE-
SECTION_INFO to store the information into the symbol_ref, and then check for it
here. When you see a const, you will have to look inside it to find the symbol_ref
in order to determine the section. See Section 17.21 [Assembler Format], page 497.
Some ports are still using a deprecated legacy substitute for this hook, the GO_IF-
LEGITIMATE_ADDRESS macro. This macro has this syntax:
#define GO_IF_LEGITIMATE_ADDRESS (mode, x, label)
and should goto label if the address x is a valid address on the target machine for a memory operand of mode mode.

Compiler source files that want to use the strict variant of this macro define the macro REG_OK STRICT. You should use an #ifdef REG_OK STRICT conditional to define the strict variant in that case and the non-strict variant otherwise.

Using the hook is usually simpler because it limits the number of files that are re-compiled when changes are made.

TARGET_MEM_CONSTRAINT [Macro]
A single character to be used instead of the default 'm' character for general memory addresses. This defines the constraint letter which matches the memory addresses accepted by TARGET_LEGITIMATE_ADDRESS P. Define this macro if you want to support new address formats in your back end without changing the semantics of the 'm' constraint. This is necessary in order to preserve functionality of inline assembly constructs using the 'm' constraint.

FIND_BASE_TERM (x) [Macro]
A C expression to determine the base term of address x, or to provide a simplified version of x from which 'alias.c' can easily find the base term. This macro is used in only two places: find_base_value and find_base_term in 'alias.c'.

It is always safe for this macro to not be defined. It exists so that alias analysis can understand machine-dependent addresses.

The typical use of this macro is to handle addresses containing a label_ref or symbol_ref within an UNSPEC.

rtx TARGET_LEGITIMIZE_ADDRESS (rtx x, rtx oldx, enum machine_mode mode) [Target Hook]
This hook is given an invalid memory address x for an operand of mode mode and should try to return a valid memory address.

x will always be the result of a call to break_out_memory_refs, and oldx will be the operand that was given to that function to produce x.

The code of the hook should not alter the substructure of x. If it transforms x into a more legitimate form, it should return the new x.

It is not necessary for this hook to come up with a legitimate address. The compiler has standard ways of doing so in all cases. In fact, it is safe to omit this hook or make it return x if it cannot find a valid way to legitimize the address. But often a machine-dependent strategy can generate better code.

LEGITIMIZE_RELOAD_ADDRESS (x, mode, opnum, type, ind_levels, win) [Macro]
A C compound statement that attempts to replace x, which is an address that needs reloading, with a valid memory address for an operand of mode mode. win will be a C statement label elsewhere in the code. It is not necessary to define this macro, but it might be useful for performance reasons.

For example, on the i386, it is sometimes possible to use a single reload register instead of two by reloading a sum of two pseudo registers into a register. On the other hand,
for number of RISC processors offsets are limited so that often an intermediate address needs to be generated in order to address a stack slot. By defining `LEGITIMIZE_RELOAD_ADDRESS` appropriately, the intermediate addresses generated for adjacent some stack slots can be made identical, and thus be shared.

*Note:* This macro should be used with caution. It is necessary to know something of how reload works in order to effectively use this, and it is quite easy to produce macros that build in too much knowledge of reload internals.

*Note:* This macro must be able to reload an address created by a previous invocation of this macro. If it fails to handle such addresses then the compiler may generate incorrect code or abort.

The macro definition should use `push_reload` to indicate parts that need reloading; `opnum`, `type` and `ind_levels` are usually suitable to be passed unaltered to `push_reload`.

The code generated by this macro must not alter the substructure of `x`. If it transforms `x` into a more legitimate form, it should assign `x` (which will always be a C variable) a new value. This also applies to parts that you change indirectly by calling `push_reload`.

The macro definition may use `strict_memory_address_p` to test if the address has become legitimate.

If you want to change only a part of `x`, one standard way of doing this is to use `copy_rtx`. Note, however, that it unshares only a single level of rtl. Thus, if the part to be changed is not at the top level, you’ll need to replace first the top level.

It is not necessary for this macro to come up with a legitimate address; but often a machine-dependent strategy can generate better code.

```c
bool TARGET_MODE_DEPENDENT_ADDRESS_P (const rtx addr) [Target Hook]
This hook returns true if memory address addr can have different meanings depending on the machine mode of the memory reference it is used for or if the address is valid for some modes but not others.
Autoincrement and autodecrement addresses typically have mode-dependent effects because the amount of the increment or decrement is the size of the operand being addressed. Some machines have other mode-dependent addresses. Many RISC machines have no mode-dependent addresses.
You may assume that addr is a valid address for the machine.
The default version of this hook returns false.
```

```c
GO_IF_MODE_DEPENDENT_ADDRESS (addr, label) [Macro]
A C statement or compound statement with a conditional goto label; executed if memory address x (an RTX) can have different meanings depending on the machine mode of the memory reference it is used for or if the address is valid for some modes but not others.
Autoincrement and autodecrement addresses typically have mode-dependent effects because the amount of the increment or decrement is the size of the operand being addressed. Some machines have other mode-dependent addresses. Many RISC machines have no mode-dependent addresses.
```
You may assume that `addr` is a valid address for the machine.

These are obsolete macros, replaced by the `TARGET_MODE_DEPENDENT_ADDRESS_P` target hook.

**LEGITIMATE_CONSTANT_P (x)**  
A C expression that is nonzero if `x` is a legitimate constant for an immediate operand on the target machine. You can assume that `x` satisfies `CONSTANT_P`, so you need not check this. In fact, ‘1’ is a suitable definition for this macro on machines where anything `CONSTANT_P` is valid.

**rtx TARGET_DELEGITIMIZE_ADDRESS (rtx x)**  
This hook is used to undo the possibly obfuscating effects of the `LEGITIMIZE_ADDRESS` and `LEGITIMIZE_RELOAD_ADDRESS` target macros. Some backend implementations of these macros wrap symbol references inside an `UNSPEC` rtx to represent PIC or similar addressing modes. This target hook allows GCC's optimizers to understand the semantics of these opaque `UNSPECs` by converting them back into their original form.

**bool TARGET_CANNOT_FORCE_CONST_MEM (rtx x)**  
This hook should return true if `x` is of a form that cannot (or should not) be spilled to the constant pool. The default version of this hook returns false.

The primary reason to define this hook is to prevent reload from deciding that a non-legitimate constant would be better reloaded from the constant pool instead of spilling and reloading a register holding the constant. This restriction is often true of addresses of TLS symbols for various targets.

**bool TARGET_USE_BLOCKS_FOR_CONSTANT_P (enum machine_mode mode, const rtx x)**  
This hook should return true if pool entries for constant `x` can be placed in an `object_block` structure. `mode` is the mode of `x`.

The default version returns false for all constants.

**tree TARGET_BUILTIN_RECIPROCAL (unsigned fn, bool md_fn, bool sqrt)**  
This hook should return the DECL of a function that implements reciprocal of the builtin function with builtin function code `fn`, or `NULL_TREE` if such a function is not available. `md_fn` is true when `fn` is a code of a machine-dependent builtin function. When `sqrt` is true, additional optimizations that apply only to the reciprocal of a square root function are performed, and only reciprocals of `sqrt` function are valid.

**tree TARGET VECTORIZE_BUILTIN MASK FOR LOAD (void)**  
This hook should return the DECL of a function `f` that given an address `addr` as an argument returns a mask `m` that can be used to extract from two vectors the relevant data that resides in `addr` in case `addr` is not properly aligned.

The autovectorizer, when vectorizing a load operation from an address `addr` that may be unaligned, will generate two vector loads from the two aligned addresses around `addr`. It then generates a `REALIGN_LOAD` operation to extract the relevant data from the two loaded vectors. The first two arguments to `REALIGN_LOAD`, `v1` and `v2`, are
the two vectors, each of size $V_S$, and the third argument, $OFF$, defines how the data will be extracted from these two vectors: if $OFF$ is 0, then the returned vector is $v_2$; otherwise, the returned vector is composed from the last $V_S - OFF$ elements of $v_1$ concatenated to the first $OFF$ elements of $v_2$.

If this hook is defined, the autovectorizer will generate a call to $f$ (using the DECL tree that this hook returns) and will use the return value of $f$ as the argument $OFF$ to `REALIGN_LOAD`. Therefore, the mask $m$ returned by $f$ should comply with the semantics expected by `REALIGN_LOAD` described above. If this hook is not defined, then $addr$ will be used as the argument $OFF$ to `REALIGN_LOAD`, in which case the low $\log_2(V_S) - 1$ bits of $addr$ will be considered.

```c
TREE TARGET_VECTORIZE_BUILTIN_MUL_WIDEN_EVEN (TREE x)  
```

[Target Hook]

This hook should return the DECL of a function $f$ that implements widening multiplication of the even elements of two input vectors of type $x$.

If this hook is defined, the autovectorizer will use it along with the `TARGET_VECTORIZE_BUILTIN_MUL_WIDEN_ODD` target hook when vectorizing widening multiplication in cases that the order of the results does not have to be preserved (e.g. used only by a reduction computation). Otherwise, the `widen_mult_hi/lo` idioms will be used.

```c
TREE TARGET_VECTORIZE_BUILTIN_MUL_WIDEN_ODD (TREE x)  
```

[Target Hook]

This hook should return the DECL of a function $f$ that implements widening multiplication of the odd elements of two input vectors of type $x$.

If this hook is defined, the autovectorizer will use it along with the `TARGET_VECTORIZE_BUILTIN_MUL_WIDEN_EVEN` target hook when vectorizing widening multiplication in cases that the order of the results does not have to be preserved (e.g. used only by a reduction computation). Otherwise, the `widen_mult_hi/lo` idioms will be used.

```c
INT TARGET_VECTORIZE_BUILTIN_VECTORIZATION_COST (ENUM vect_cost_for_stmt type_of_cost, TREE vectype, INT misalign)  
```

[Target Hook]

Returns cost of different scalar or vector statements for vectorization cost model. For vector memory operations the cost may depend on type ($vectype$) and misalignment value ($misalign$).

```c
BOOL TARGET_VECTORIZE_VECTOR_ALIGNMENT_REACHABLE  
```

(const TREE type, BOOL is_packed)

Return true if vector alignment is reachable (by peeling N iterations) for the given type.

```c
TREE TARGET_VECTORIZE_BUILTIN_VEC_PERM (TREE type, TREE *mask_element_type)  
```

[Target Hook]

Target builtin that implements vector permute.

```c
BOOL TARGET_VECTORIZE_BUILTIN_VEC_PERM_OK (TREE vec_type,  
```

(TREE mask)

Return true if a vector created for `builtin_vec_perm` is valid.
tree TARGET_VECTORIZE_BUILTIN_CONVERSION (unsigned code, [Target Hook]
  tree dest_type, tree src_type)
This hook should return the DECL of a function that implements conversion of the
input vector of type src_type to type dest_type. The value of code is one of the
enumerators in enum tree_code and specifies how the conversion is to be applied
(truncation, rounding, etc.).
If this hook is defined, the autovectorizer will use the TARGET_VECTORIZE_BUILTIN_-
CONVERSION target hook when vectorizing conversion. Otherwise, it will return NULL_-
TREE.

tree TARGET_VECTORIZE_BUILTIN_VECTORIZED_FUNCTION (tree [Target Hook]
  fndecl, tree vec_type_out, tree vec_type_in)
This hook should return the decl of a function that implements the vectorized variant
of the builtin function with builtin function code code or NULL_TREE if such a function
is not available. The value of fndecl is the builtin function declaration. The return
type of the vectorized function shall be of vector type vec_type_out and the argument
types should be vec_type_in.

bool TARGET_VECTORIZE_SUPPORT_VECTOR_MISALIGNMENT (enum [Target Hook]
  machine_mode mode, const tree type, int misalignment, bool is_packed)
This hook should return true if the target supports misaligned vector store/load of a
specific factor denoted in the misalignment parameter. The vector store/load should
be of machine mode mode and the elements in the vectors should be of type type.
is_packed parameter is true if the memory access is defined in a packed struct.

enum machine_mode TARGET_VECTORIZE_PREFERRED_SIMD_MODE [Target Hook]
  (enum machine_mode mode)
This hook should return the preferred mode for vectorizing scalar mode mode. The
default is equal to word_mode, because the vectorizer can do some transformations
even in absence of specialized SIMD hardware.

unsigned int TARGET_VECTORIZE_AUTOVECTORIZE_VECTOR_SIZES (void) [Target Hook]
This hook should return a mask of sizes that should be iterated over after trying
to autovectorize using the vector size derived from the mode returned by TARGET_-
VECTORIZE_PREFERRED_SIMD_MODE. The default is zero which means to not iterate
over other vector sizes.

17.15 Anchored Addresses
GCC usually addresses every static object as a separate entity. For example, if we have:

```c
static int a, b, c;
int foo (void) { return a + b + c; }
```

the code for foo will usually calculate three separate symbolic addresses: those of a, b
and c. On some targets, it would be better to calculate just one symbolic address and access
the three variables relative to it. The equivalent pseudocode would be something like:

```c
int foo (void)
{
  register int *xr = &x;
```

return \texttt{xr[\&a - \&x] + xr[\&b - \&x] + xr[\&c - \&x];}

(which isn’t valid C). We refer to shared addresses like \texttt{x} as “section anchors”. Their use
is controlled by ‘-fsection-anchors’.

The hooks below describe the target properties that GCC needs to know in order to
make effective use of section anchors. It won’t use section anchors at all unless either
\texttt{TARGET_MIN_ANCHOR_OFFSET} or \texttt{TARGET_MAX_ANCHOR_OFFSET} is set to a nonzero value.

\texttt{HOST_WIDE_INT TARGET_MIN_ANCHOR_OFFSET} [Target Hook]
The minimum offset that should be applied to a section anchor. On most targets, it
should be the smallest offset that can be applied to a base register while still giving
a legitimate address for every mode. The default value is 0.

\texttt{HOST_WIDE_INT TARGET_MAX_ANCHOR_OFFSET} [Target Hook]
Like \texttt{TARGET_MIN_ANCHOR_OFFSET}, but the maximum (inclusive) offset that should be
applied to section anchors. The default value is 0.

\texttt{void TARGET_ASM_OUTPUT_ANCHOR (rtx x)} [Target Hook]
Write the assembly code to define section anchor \texttt{x}, which is a \texttt{SYMBOL_REF}
for which ‘\texttt{SYMBOL_REF_ANCHOR_P (x)}’ is true. The hook is called with the assembly output
position set to the beginning of \texttt{SYMBOL_REF_BLOCK (x)}.

If \texttt{ASM_OUTPUT_DEF} is available, the hook’s default definition uses it to define the sym-
bol as ‘. + SYMBOL_REF_BLOCK_OFFSET (x)’. If \texttt{ASM_OUTPUT_DEF} is not available, the
hook’s default definition is \texttt{NULL}, which disables the use of section anchors altogether.

\texttt{bool TARGET_USE_ANCHORS_FOR_SYMBOL_P (const_rtx x)} [Target Hook]
Return true if GCC should attempt to use anchors to access \texttt{SYMBOL_REF x}. You can
assume ‘\texttt{SYMBOL_REF_HAS_BLOCK_INFO_P (x)}’ and ‘!\texttt{SYMBOL_REF_ANCHOR_P (x)}’.

The default version is correct for most targets, but you might need to intercept this
hook to handle things like target-specific attributes or target-specific sections.

17.16 Condition Code Status

The macros in this section can be split in two families, according to the two ways of repre-
senting condition codes in GCC.

The first representation is the so called (cc0) representation (see Section 16.12 [Jump
Patterns], page 357), where all instructions can have an implicit clobber of the condition
codes. The second is the condition code register representation, which provides better
schedulability for architectures that do have a condition code register, but on which most
instructions do not affect it. The latter category includes most RISC machines.

The implicit clobbering poses a strong restriction on the placement of the definition and
use of the condition code, which need to be in adjacent insns for machines using (cc0).
This can prevent important optimizations on some machines. For example, on the IBM
RS/6000, there is a delay for taken branches unless the condition code register is set three
instructions earlier than the conditional branch. The instruction scheduler cannot perform
this optimization if it is not permitted to separate the definition and use of the condition
code register.
For this reason, it is possible and suggested to use a register to represent the condition code for new ports. If there is a specific condition code register in the machine, use a hard register. If the condition code or comparison result can be placed in any general register, or if there are multiple condition registers, use a pseudo register. Registers used to store the condition code value will usually have a mode that is in class MODE_CC.

Alternatively, you can use BImode if the comparison operator is specified already in the compare instruction. In this case, you are not interested in most macros in this section.

### 17.16.1 Representation of condition codes using (cc0)

The file ‘conditions.h’ defines a variable cc_status to describe how the condition code was computed (in case the interpretation of the condition code depends on the instruction that it was set by). This variable contains the RTL expressions on which the condition code is currently based, and several standard flags.

Sometimes additional machine-specific flags must be defined in the machine description header file. It can also add additional machine-specific information by defining CC_STATUS_MDEP.

**CC_STATUS_MDEP**  
C code for a data type which is used for declaring the mdep component of cc_status. It defaults to int.

This macro is not used on machines that do not use cc0.

**CC_STATUS_MDEP_INIT**  
A C expression to initialize the mdep field to “empty”. The default definition does nothing, since most machines don’t use the field anyway. If you want to use the field, you should probably define this macro to initialize it.

This macro is not used on machines that do not use cc0.

**NOTICE_UPDATE_CC (exp, insn)**  
A C compound statement to set the components of cc_status appropriately for an insn insn whose body is exp. It is this macro’s responsibility to recognize insns that set the condition code as a byproduct of other activity as well as those that explicitly set (cc0).

This macro is not used on machines that do not use cc0.

If there are insns that do not set the condition code but do alter other machine registers, this macro must check to see whether they invalidate the expressions that the condition code is recorded as reflecting. For example, on the 68000, insns that store in address registers do not set the condition code, which means that usually NOTICE_UPDATE_CC can leave cc_status unaltered for such insns. But suppose that the previous insn set the condition code based on location ‘a4@{(102)}’ and the current insn stores a new value in ‘a4’. Although the condition code is not changed by this, it will no longer be true that it reflects the contents of ‘a4@{(102)}’. Therefore, NOTICE_UPDATE_CC must alter cc_status in this case to say that nothing is known about the condition code value.

The definition of NOTICE_UPDATE_CC must be prepared to deal with the results of peephole optimization: insns whose patterns are parallel RTXs containing various
**reg, mem** or constants which are just the operands. The RTL structure of these insns is not sufficient to indicate what the insns actually do. What **NOTICE_UPDATE_CC** should do when it sees one is just to run **CC_STATUS_INIT**.

A possible definition of **NOTICE_UPDATE_CC** is to call a function that looks at an attribute (see Section 16.19 [Insn Attributes], page 371) named, for example, ‘cc’. This avoids having detailed information about patterns in two places, the ‘md’ file and in **NOTICE_UPDATE_CC**.

### 17.16.2 Representation of condition codes using registers

**SELECT_CC_MODE** *(op, x, y)*  

On many machines, the condition code may be produced by other instructions than compares, for example the branch can use directly the condition code set by a subtract instruction. However, on some machines when the condition code is set this way some bits (such as the overflow bit) are not set in the same way as a test instruction, so that a different branch instruction must be used for some conditional branches. When this happens, use the machine mode of the condition code register to record different formats of the condition code register. Modes can also be used to record which compare instruction (e.g. a signed or an unsigned comparison) produced the condition codes.

If other modes than **CCmode** are required, add them to ‘machine-modes.def’ and define **SELECT_CC_MODE** to choose a mode given an operand of a compare. This is needed because the modes have to be chosen not only during RTL generation but also, for example, by instruction combination. The result of **SELECT_CC_MODE** should be consistent with the mode used in the patterns; for example to support the case of the add on the SPARC discussed above, we have the pattern

```c
(define_insn ""
  [(set (reg:CC_NOOV 0)
      (compare:CC_NOOV
        (plus:SI (match_operand:SI 0 "register_operand" "%r")
          (match_operand:SI 1 "arith_operand" "rI")))
      (const_int 0))]
""
"
```

together with a **SELECT_CC_MODE** that returns **CC_NOOVmode** for comparisons whose argument is a plus:

```c
#define SELECT_CC_MODE(OP,X,Y) \
  (GET_MODE_CLASS (GET_MODE (X)) == MODE_FLOAT \ 
   ? ((OP == EQ || OP == NE) ? CCFPmode : CCFPEmode) \ 
    : ((GET_CODE (X) == PLUS || GET_CODE (X) == MINUS \ 
       || GET_CODE (X) == NEG) \ 
       ? CC_NOOVmode : CCmode))
```

Another reason to use modes is to retain information on which operands were used by the comparison; see **REVERSIBLE_CC_MODE** later in this section.

You should define this macro if and only if you define extra CC modes in ‘machine-modes.def’.

**CANONICALIZE_COMPARISON** *(code, op0, op1)*  

On some machines not all possible comparisons are defined, but you can convert an invalid comparison into a valid one. For example, the Alpha does not have a **GT**
comparison, but you can use an LT comparison instead and swap the order of the operands.

On such machines, define this macro to be a C statement to do any required conversions. `code` is the initial comparison code and `op0` and `op1` are the left and right operands of the comparison, respectively. You should modify `code`, `op0`, and `op1` as required.

GCC will not assume that the comparison resulting from this macro is valid but will see if the resulting insn matches a pattern in the `md` file.

You need not define this macro if it would never change the comparison code or operands.

**REVERSIBLE_CC_MODE (mode)**  
A C expression whose value is one if it is always safe to reverse a comparison whose mode is `mode`. If `SELECT_CC_MODE` can ever return `mode` for a floating-point inequality comparison, then `REVERSIBLE_CC_MODE (mode)` must be zero.

You need not define this macro if it would always return zero or if the floating-point format is anything other than `IEEE_FLOAT_FORMAT`. For example, here is the definition used on the SPARC, where floating-point inequality comparisons are always given `CCFPEmode`:

```c
#define REVERSIBLE_CC_MODE(MODE) ((MODE) != CCFPEmode)
```

**REVERSE_CONDITION (code, mode)**  
A C expression whose value is reversed condition code of the `code` for comparison done in CC `mode`. The macro is used only in case `REVERSIBLE_CC_MODE (mode)` is nonzero. Define this macro in case machine has some non-standard way how to reverse certain conditionals. For instance in case all floating point conditions are non-trapping, compiler may freely convert unordered compares to ordered one. Then definition may look like:

```c
#define REVERSE_CONDITION(CODE, MODE) 
  ((MODE) != CCFPEmode ? reverse_condition (CODE) \
   : reverse_condition_maybe_unordered (CODE))
```

**bool TARGET_FIXED_CONDITION_CODE_REGS (unsigned int *p1, unsigned int *p2)**  
On targets which do not use (cc0), and which use a hard register rather than a pseudo-register to hold condition codes, the regular CSE passes are often not able to identify cases in which the hard register is set to a common value. Use this hook to enable a small pass which optimizes such cases. This hook should return true to enable this pass, and it should set the integers to which its arguments point to the hard register numbers used for condition codes. When there is only one such register, as is true on most systems, the integer pointed to by `p2` should be set to `INVALID_REGNUM`.

The default version of this hook returns false.

**enum machine_mode TARGET_CC_MODES_COMPATIBLE (enum machine_mode m1, enum machine_mode m2)**  
On targets which use multiple condition code modes in class `MODE_CC`, it is sometimes the case that a comparison can be validly done in more than one mode. On such a
system, define this target hook to take two mode arguments and to return a mode in which both comparisons may be validly done. If there is no such mode, return VOIDmode.

The default version of this hook checks whether the modes are the same. If they are, it returns that mode. If they are different, it returns VOIDmode.

17.16.3 Macros to control conditional execution

There is one macro that may need to be defined for targets supporting conditional execution, independent of how they represent conditional branches.

REVERSE_CONDEXEC_PREDICATES_P (op1, op2) [Macro]
A C expression that returns true if the conditional execution predicate op1, a comparison operation, is the inverse of op2 and vice versa. Define this to return 0 if the target has conditional execution predicates that cannot be reversed safely. There is no need to validate that the arguments of op1 and op2 are the same, this is done separately. If no expansion is specified, this macro is defined as follows:

#define REVERSE_CONDEXEC_PREDICATES_P (x, y) \  (GET_CODE ((x)) == reversed_comparison_code ((y), NULL))

17.17 Describing Relative Costs of Operations

These macros let you describe the relative speed of various operations on the target machine.

REGISTER_MOVE_COST (mode, from, to) [Macro]
A C expression for the cost of moving data of mode mode from a register in class from to one in class to. The classes are expressed using the enumeration values such as GENERAL_REGS. A value of 2 is the default; other values are interpreted relative to that.

It is not required that the cost always equal 2 when from is the same as to; on some machines it is expensive to move between registers if they are not general registers.

If reload sees an insn consisting of a single set between two hard registers, and if REGISTER_MOVE_COST applied to their classes returns a value of 2, reload does not check to ensure that the constraints of the insn are met. Setting a cost of other than 2 will allow reload to verify that the constraints are met. You should do this if the ‘movm’ pattern’s constraints do not allow such copying.

These macros are obsolete, new ports should use the target hook TARGET_REGISTER_MOVE_COST instead.

int TARGET_REGISTER_MOVE_COST (enum machine_mode mode, \  reg_class_t from, reg_class_t to) [Target Hook]
This target hook should return the cost of moving data of mode mode from a register in class from to one in class to. The classes are expressed using the enumeration values such as GENERAL_REGS. A value of 2 is the default; other values are interpreted relative to that.

It is not required that the cost always equal 2 when from is the same as to; on some machines it is expensive to move between registers if they are not general registers.
If reload sees an insn consisting of a single set between two hard registers, and if \texttt{TARGET\_REGISTER\_MOVE\_COST} applied to their classes returns a value of 2, reload does not check to ensure that the constraints of the insn are met. Setting a cost of other than 2 will allow reload to verify that the constraints are met. You should do this if the ‘\texttt{movm}’ pattern’s constraints do not allow such copying.

The default version of this function returns 2.

\textbf{MEMORY\_MOVE\_COST (mode, class, in)}

A C expression for the cost of moving data of mode \texttt{mode} between a register of class \texttt{class} and memory; \texttt{in} is zero if the value is to be written to memory, nonzero if it is to be read in. This cost is relative to those in \texttt{REGISTER\_MOVE\_COST}. If moving between registers and memory is more expensive than between two registers, you should define this macro to express the relative cost.

If you do not define this macro, GCC uses a default cost of 4 plus the cost of copying via a secondary reload register, if one is needed. If your machine requires a secondary reload register to copy between memory and a register of \texttt{class} but the reload mechanism is more complex than copying via an intermediate, define this macro to reflect the actual cost of the move.

GCC defines the function \texttt{memory\_move\_secondary\_cost} if secondary reloads are needed. It computes the costs due to copying via a secondary register. If your machine copies from memory using a secondary register in the conventional way but the default base value of 4 is not correct for your machine, define this macro to add some other value to the result of that function. The arguments to that function are the same as to this macro.

These macros are obsolete, new ports should use the target hook \texttt{TARGET\_MEMORY\_MOVE\_COST} instead.

\textbf{int TARGET\_MEMORY\_MOVE\_COST (enum machine\_mode mode, reg\_class\_t rclass, bool in)}

This target hook should return the cost of moving data of mode \texttt{mode} between a register of class \texttt{rclass} and memory; \texttt{in} is \texttt{false} if the value is to be written to memory, \texttt{true} if it is to be read in. This cost is relative to those in \texttt{TARGET\_REGISTER\_MOVE\_COST}. If moving between registers and memory is more expensive than between two registers, you should add this target hook to express the relative cost.

If you do not add this target hook, GCC uses a default cost of 4 plus the cost of copying via a secondary reload register, if one is needed. If your machine requires a secondary reload register to copy between memory and a register of \texttt{rclass} but the reload mechanism is more complex than copying via an intermediate, use this target hook to reflect the actual cost of the move.

GCC defines the function \texttt{memory\_move\_secondary\_cost} if secondary reloads are needed. It computes the costs due to copying via a secondary register. If your machine copies from memory using a secondary register in the conventional way but the default base value of 4 is not correct for your machine, use this target hook to add some other value to the result of that function. The arguments to that function are the same as to this target hook.
**BRANCH_COST** (*speed_p, predictable_p*)  
A C expression for the cost of a branch instruction. A value of 1 is the default; other values are interpreted relative to that. Parameter *speed_p* is true when the branch in question should be optimized for speed. When it is false, **BRANCH_COST** should return a value optimal for code size rather than performance. *predictable_p* is true for well-predicted branches. On many architectures the **BRANCH_COST** can be reduced then.

Here are additional macros which do not specify precise relative costs, but only that certain actions are more expensive than GCC would ordinarily expect.

**SLOW_BYTE_ACCESS**  
Define this macro as a C expression which is nonzero if accessing less than a word of memory (i.e. a *char* or a *short*) is no faster than accessing a word of memory, i.e., if such access require more than one instruction or if there is no difference in cost between byte and (aligned) word loads.

When this macro is not defined, the compiler will access a field by finding the smallest containing object; when it is defined, a fullword load will be used if alignment permits. Unless bytes accesses are faster than word accesses, using word accesses is preferable since it may eliminate subsequent memory access if subsequent accesses occur to other fields in the same word of the structure, but to different bytes.

**SLOW_UNALIGNED_ACCESS** (*mode, alignment*)  
Define this macro to be the value 1 if memory accesses described by the *mode* and *alignment* parameters have a cost many times greater than aligned accesses, for example if they are emulated in a trap handler.

When this macro is nonzero, the compiler will act as if **STRICT_ALIGNMENT** were nonzero when generating code for block moves. This can cause significantly more instructions to be produced. Therefore, do not set this macro nonzero if unaligned accesses only add a cycle or two to the time for a memory access.

If the value of this macro is always zero, it need not be defined. If this macro is defined, it should produce a nonzero value when **STRICT_ALIGNMENT** is nonzero.

**MOVE_RATIO** (*speed*)  
The threshold of number of scalar memory-to-memory move insns, below which a sequence of insns should be generated instead of a string move insn or a library call. Increasing the value will always make code faster, but eventually incurs high cost in increased code size.

Note that on machines where the corresponding move insn is a **define_expand** that emits a sequence of insns, this macro counts the number of such sequences.

The parameter *speed* is true if the code is currently being optimized for speed rather than size.

If you don’t define this, a reasonable default is used.

**MOVE_BYPieces_P** (*size, alignment*)  
A C expression used to determine whether move_by_pieces will be used to copy a chunk of memory, or whether some other block move mechanism will be used. Defaults to 1 if move_by_pieces_ninsns returns less than **MOVE_RATIO**.
MOVE_MAX_PIECES [Macro]
A C expression used by move_by_pieces to determine the largest unit a load or store used to copy memory is. Defaults to MOVE_MAX.

CLEAR_RATIO (speed) [Macro]
The threshold of number of scalar move insns, below which a sequence of insns should be generated to clear memory instead of a string clear insn or a library call. Increasing the value will always make code faster, but eventually incurs high cost in increased code size.
The parameter speed is true if the code is currently being optimized for speed rather than size.
If you don’t define this, a reasonable default is used.

CLEAR_BY_PIECES_P (size, alignment) [Macro]
A C expression used to determine whether clear_by_pieces will be used to clear a chunk of memory, or whether some other block clear mechanism will be used. Defaults to 1 if move_by_pieces_ninsns returns less than CLEAR_RATIO.

SET_RATIO (speed) [Macro]
The threshold of number of scalar move insns, below which a sequence of insns should be generated to set memory to a constant value, instead of a block set insn or a library call. Increasing the value will always make code faster, but eventually incurs high cost in increased code size.
The parameter speed is true if the code is currently being optimized for speed rather than size.
If you don’t define this, it defaults to the value of MOVE_RATIO.

SET_BY_PIECES_P (size, alignment) [Macro]
A C expression used to determine whether store_by_pieces will be used to set a chunk of memory to a constant value, or whether some other mechanism will be used. Used by __builtin_memset when storing values other than constant zero. Defaults to 1 if move_by_pieces_ninsns returns less than SET_RATIO.

STORE_BY_PIECES_P (size, alignment) [Macro]
A C expression used to determine whether store_by_pieces will be used to set a chunk of memory to a constant string value, or whether some other mechanism will be used. Used by __builtin_strcpy when called with a constant source string. Defaults to 1 if move_by_pieces_ninsns returns less than MOVE_RATIO.

USE_LOAD_POST_INCREMENT (mode) [Macro]
A C expression used to determine whether a load postincrement is a good thing to use for a given mode. Defaults to the value of HAVE_POST_INCREMENT.

USE_LOAD_POST_DECREMENT (mode) [Macro]
A C expression used to determine whether a load postdecrement is a good thing to use for a given mode. Defaults to the value of HAVE_POST_DECREMENT.

USE_LOAD_PRE_INCREMENT (mode) [Macro]
A C expression used to determine whether a load preincrement is a good thing to use for a given mode. Defaults to the value of HAVE_PRE_INCREMENT.
USE_LOAD_PRE_DECREMENT (mode) [Macro]
A C expression used to determine whether a load predecrement is a good thing to use for a given mode. Defaults to the value of HAVE_PRE_DECREMENT.

USE_STORE_POST_INCREMENT (mode) [Macro]
A C expression used to determine whether a store postincrement is a good thing to use for a given mode. Defaults to the value of HAVE_POST_INCREMENT.

USE_STORE_POST_DECREMENT (mode) [Macro]
A C expression used to determine whether a store postdecrement is a good thing to use for a given mode. Defaults to the value of HAVE_POST_DECREMENT.

USE_STORE_PRE_INCREMENT (mode) [Macro]
This macro is used to determine whether a store preincrement is a good thing to use for a given mode. Defaults to the value of HAVE_PRE_INCREMENT.

USE_STORE_PRE_DECREMENT (mode) [Macro]
This macro is used to determine whether a store predecrement is a good thing to use for a given mode. Defaults to the value of HAVE_PRE_DECREMENT.

NO_FUNCTION_CSE [Macro]
Define this macro if it is as good or better to call a constant function address than to call an address kept in a register.

RANGE_TEST_NON_SHORT_CIRCUIT [Macro]
Define this macro if a non-short-circuit operation produced by ‘fold_range_test ()’ is optimal. This macro defaults to true if BRANCH_COST is greater than or equal to the value 2.

bool TARGET_RTX_COSTS (rtx x, int code, int outer_code, int *total, bool speed) [Target Hook]
This target hook describes the relative costs of RTL expressions.

The cost may depend on the precise form of the expression, which is available for examination in x, and the rtx code of the expression in which it is contained, found in outer_code. code is the expression code—redundant, since it can be obtained with GET_CODE (x).

In implementing this hook, *total contains a default estimate for the cost of the expression. The hook should modify this value as necessary. Traditionally, the default costs are COSTS_N_INSNS (5) for multiplications, COSTS_N_INSNS (7) for division and modulus operations, and COSTS_N_INSNS (1) for all other operations.

When optimizing for code size, i.e. when speed is false, this target hook should be used to estimate the relative size cost of an expression, again relative to COSTS_N_INSNS.

The hook returns true when all subexpressions of x have been processed, and false when rtx_cost should recurse.
int TARGET_ADDRESS_COST (rtx address, bool speed)   [Target Hook]
This hook computes the cost of an addressing mode that contains address. If not defined, the cost is computed from the address expression and the TARGET_RTX_COST hook.

For most CISC machines, the default cost is a good approximation of the true cost of the addressing mode. However, on RISC machines, all instructions normally have the same length and execution time. Hence all addresses will have equal costs.

In cases where more than one form of an address is known, the form with the lowest cost will be used. If multiple forms have the same, lowest, cost, the one that is the most complex will be used.

For example, suppose an address that is equal to the sum of a register and a constant is used twice in the same basic block. When this macro is not defined, the address will be computed in a register and memory references will be indirect through that register. On machines where the cost of the addressing mode containing the sum is no higher than that of a simple indirect reference, this will produce an additional instruction and possibly require an additional register. Proper specification of this macro eliminates this overhead for such machines.

This hook is never called with an invalid address.

On machines where an address involving more than one register is as cheap as an address computation involving only one register, defining TARGET_ADDRESS_COST to reflect this can cause two registers to be live over a region of code where only one would have been if TARGET_ADDRESS_COST were not defined in that manner. This effect should be considered in the definition of this macro. Equivalent costs should probably only be given to addresses with different numbers of registers on machines with lots of registers.

17.18 Adjusting the Instruction Scheduler

The instruction scheduler may need a fair amount of machine-specific adjustment in order to produce good code. GCC provides several target hooks for this purpose. It is usually enough to define just a few of them: try the first ones in this list first.

int TARGET_SCHED_ISSUE_RATE (void)   [Target Hook]
This hook returns the maximum number of instructions that can ever issue at the same time on the target machine. The default is one. Although the insn scheduler can define itself the possibility of issue an insn on the same cycle, the value can serve as an additional constraint to issue insns on the same simulated processor cycle (see hooks ‘TARGET_SCHED_REORDER’ and ‘TARGET_SCHED_REORDER2’). This value must be constant over the entire compilation. If you need it to vary depending on what the instructions are, you must use ‘TARGET_SCHED_VARIABLE_ISSUE’.

int TARGET_SCHED_VARIABLE_ISSUE (FILE *file, int verbose, rtx insn, int more)   [Target Hook]
This hook is executed by the scheduler after it has scheduled an insn from the ready list. It should return the number of insns which can still be issued in the current cycle. The default is ‘more - 1’ for insns other than CLOBBER and USE, which normally are not counted against the issue rate. You should define this hook if some insns take
more machine resources than others, so that fewer insns can follow them in the same

cycle. file is either a null pointer, or a stdio stream to write any debug output to.
verbose is the verbose level provided by `-fsched-verbose-n`. insn is the instruction
that was scheduled.

int TARGET_SCHED_ADJUST_COST (rtx insn, rtx link, rtx dep_insn, int cost)  [Target Hook]

This function corrects the value of cost based on the relationship between insn and
dep_insn through the dependence link. It should return the new value. The default is
to make no adjustment to cost. This can be used for example to specify to the sched-
uler using the traditional pipeline description that an output- or anti-dependence does
not incur the same cost as a data-dependence. If the scheduler using the automaton
based pipeline description, the cost of anti-dependence is zero and the cost of output-
dependence is maximum of one and the difference of latency times of the first and the
second insns. If these values are not acceptable, you could use the hook to modify
them too. See also see Section 16.19.8 [Processor pipeline description], page 380.

int TARGET_SCHED_ADJUST_PRIORITY (rtx insn, int priority)  [Target Hook]

This hook adjusts the integer scheduling priority priority of insn. It should return
the new priority. Increase the priority to execute insn earlier, reduce the priority to
execute insn later. Do not define this hook if you do not need to adjust the scheduling
priorities of insns.

int TARGET_SCHED_REORDER (FILE *file, int verbose, rtx *ready, int *n_readyp, int clock)  [Target Hook]

This hook is executed by the scheduler after it has scheduled the ready list, to allow
the machine description to reorder it (for example to combine two small instructions
together on ‘VLIW’ machines). file is either a null pointer, or a stdio stream to write
any debug output to. verbose is the verbose level provided by ‘-fsched-verbose-n’. 
ready is a pointer to the ready list of instructions that are ready to be scheduled. 
n_readyp is a pointer to the number of elements in the ready list. The scheduler
reads the ready list in reverse order, starting with ready[*n_readyp – 1] and going to
ready[0]. clock is the timer tick of the scheduler. You may modify the ready list and
the number of ready insns. The return value is the number of insns that can issue
this cycle; normally this is just issue_rate. See also ‘TARGET_SCHED_REORDER2’.

int TARGET_SCHED_REORDER2 (FILE *file, int verbose, rtx *ready, int *n_readyp, int clock)  [Target Hook]

Like ‘TARGET_SCHED_REORDER’, but called at a different time. That function is called
whenever the scheduler starts a new cycle. This one is called once per iteration over
a cycle, immediately after ‘TARGET_SCHED_VARIABLE_ISSUE’; it can reorder the ready
list and return the number of insns to be scheduled in the same cycle. Defining this
hook can be useful if there are frequent situations where scheduling one insn causes
other insns to become ready in the same cycle. These other insns can then be taken
into account properly.
void TARGET_SCHED_DEPENDENCIES_EVALUATION_HOOK (rtx head, rtx tail)  
This hook is called after evaluation forward dependencies of insns in chain given by  
two parameter values (head and tail correspondingly) but before insns scheduling of  
the insn chain. For example, it can be used for better insn classification if it requires  
analysis of dependencies. This hook can use backward and forward dependencies of  
the insn scheduler because they are already calculated.

void TARGET_SCHED_INIT (FILE *file, int verbose, int max_ready)  
This hook is executed by the scheduler at the beginning of each block of instructions  
that are to be scheduled. file is either a null pointer, or a stdio stream to write  
any debug output to. verbose is the verbose level provided by `-fsched-verbose-n’.  
max_ready is the maximum number of insns in the current scheduling region that can  
be live at the same time. This can be used to allocate scratch space if it is needed,  
e.g. by `TARGET_SCHED_REORDER’.

void TARGET_SCHED_FINISH (FILE *file, int verbose)  
This hook is executed by the scheduler at the end of each block of instructions that  
ard to be scheduled. It can be used to perform cleanup of any actions done by the  
other scheduling hooks. file is either a null pointer, or a stdio stream to write any  
debug output to. verbose is the verbose level provided by `-fsched-verbose-n’.

void TARGET_SCHED_INIT_GLOBAL (FILE *file, int verbose, int old_max_uid)  
This hook is executed by the scheduler after function level initializations. file is either  
a null pointer, or a stdio stream to write any debug output to. verbose is the verbose  
level provided by `-fsched-verbose-n’. old_max_uid is the maximum insn uid when  
Scheduling begins.

void TARGET_SCHED_FINISH_GLOBAL (FILE *file, int verbose)  
This is the cleanup hook corresponding to TARGET_SCHED_INIT_GLOBAL. file is either  
a null pointer, or a stdio stream to write any debug output to. verbose is the verbose  
level provided by `-fsched-verbose-n’.

rtx TARGET_SCHED_DFA_PRE_CYCLE_INSN (void)  
The hook returns an RTL insn. The automaton state used in the pipeline hazard  
recognizer is changed as if the insn were scheduled when the new simulated processor  
cycle starts. Usage of the hook may simplify the automaton pipeline description for  
some VLIW processors. If the hook is defined, it is used only for the automaton based  
pipeline description. The default is not to change the state when the new simulated  
processor cycle starts.

void TARGET_SCHED_INIT_DFA_PRE_CYCLE_INSN (void)  
The hook can be used to initialize data used by the previous hook.

rtx TARGET_SCHED_DFA_POST_CYCLE_INSN (void)  
The hook is analogous to ‘TARGET_SCHED_DFA_PRE_CYCLE_INSN’ but used to changed  
the state as if the insn were scheduled when the new simulated processor cycle finishes.
The hook is analogous to ‘TARGET_SCHED_INIT_DFA_PRE_CYCLE_INSN’ but used to initialize data used by the previous hook.

The hook to notify target that the current simulated cycle is about to finish. The hook is analogous to ‘TARGET_SCHED_DFA_PRE_CYCLE_INSN’ but used to change the state in more complicated situations - e.g., when advancing state on a single insn is not enough.

The hook to notify target that new simulated cycle has just started. The hook is analogous to ‘TARGET_SCHED_DFA_POST_CYCLE_INSN’ but used to change the state in more complicated situations - e.g., when advancing state on a single insn is not enough.

This hook controls better choosing an insn from the ready insn queue for the DFA-based insn scheduler. Usually the scheduler chooses the first insn from the queue. If the hook returns a positive value, an additional scheduler code tries all permutations of ‘TARGET_SCHED_FIRST_CYCLE_MULTIPASS_DFA_LOOKAHEAD ()’ subsequent ready insns to choose an insn whose issue will result in maximal number of issued insns on the same cycle. For the VLIW processor, the code could actually solve the problem of packing simple insns into the VLIW insn. Of course, if the rules of VLIW packing are described in the automaton.

This code also could be used for superscalar RISC processors. Let us consider a superscalar RISC processor with 3 pipelines. Some insns can be executed in pipelines A or B, some insns can be executed only in pipelines B or C, and one insn can be executed in pipeline B. The processor may issue the 1st insn into A and the 2nd one into B. In this case, the 3rd insn will wait for freeing B until the next cycle. If the scheduler issues the 3rd insn the first, the processor could issue all 3 insns per cycle.

Actually this code demonstrates advantages of the automaton based pipeline hazard recognizer. We try quickly and easy many insn schedules to choose the best one.

The default is no multipass scheduling.

This hook controls what insns from the ready insn queue will be considered for the multipass insn scheduling. If the hook returns zero for insn, the insn will be not chosen to be issued.

The default is that any ready insns can be chosen to be issued.

This hook prepares the target backend for a new round of multipass scheduling.
Chapter 17: Target Description Macros and Functions

void TARGET_SCHED_FIRST_CYCLE_MULTIPASS_ISSUE (void *data, char *ready_try, int n_ready, rtx insn, const void *prev_data)
[Target Hook]
This hook is called when multipass scheduling evaluates instruction INSN.

void TARGET_SCHED_FIRST_CYCLE_MULTIPASS_BACKTRACK (const void *data, char *ready_try, int n_ready)
[Target Hook]
This is called when multipass scheduling backtracks from evaluation of an instruction.

void TARGET_SCHED_FIRST_CYCLE_MULTIPASS_END (const void *data)
[Target Hook]
This hook notifies the target about the result of the concluded current round of multipass scheduling.

void TARGET_SCHED_FIRST_CYCLE_MULTIPASS_INIT (void *data)
[Target Hook]
This hook initializes target-specific data used in multipass scheduling.

void TARGET_SCHED_FIRST_CYCLE_MULTIPASS_FINI (void *data)
[Target Hook]
This hook finalizes target-specific data used in multipass scheduling.

int TARGET_SCHED_DFA_NEW_CYCLE (FILE *dump, int verbose, rtx insn, int last_clock, int clock, int *sort_p)
[Target Hook]
This hook is called by the insn scheduler before issuing insn on cycle clock. If the hook returns nonzero, insn is not issued on this processor cycle. Instead, the processor cycle is advanced. If *sort_p is zero, the insn ready queue is not sorted on the new cycle start as usually. dump and verbose specify the file and verbosity level to use for debugging output. last_clock and clock are, respectively, the processor cycle on which the previous insn has been issued, and the current processor cycle.

bool TARGET_SCHED_IS_COSTLY_DEPENDENCE (struct _dep *dep, int cost, int distance)
[Target Hook]
This hook is used to define which dependences are considered costly by the target, so costly that it is not advisable to schedule the insns that are involved in the dependence too close to one another. The parameters to this hook are as follows: The first parameter _dep is the dependence being evaluated. The second parameter cost is the cost of the dependence as estimated by the scheduler, and the third parameter distance is the distance in cycles between the two insns. The hook returns true if considering the distance between the two insns the dependence between them is considered costly by the target, and false otherwise.

Defining this hook can be useful in multiple-issue out-of-order machines, where (a) it’s practically hopeless to predict the actual data/resource delays, however: (b) there’s a better chance to predict the actual grouping that will be formed, and (c) correctly emulating the grouping can be very important. In such targets one may want to allow issuing dependent insns closer to one another—i.e., closer than the dependence distance; however, not in cases of “costly dependences”, which this hooks allows to define.

void TARGET_SCHED_H_I_D_EXTENDED (void)
[Target Hook]
This hook is called by the insn scheduler after emitting a new instruction to the instruction stream. The hook notifies a target backend to extend its per instruction data structures.
void * TARGET_SCHED_ALLOC_SCHED_CONTEXT (void)  
[Target Hook] 
Return a pointer to a store large enough to hold target scheduling context.

void TARGET_SCHED_INIT_SCHED_CONTEXT (void *tc, bool clean_p) 
[Target Hook] 
Initialize store pointed to by tc to hold target scheduling context. It clean_p is true 
then initialize tc as if scheduler is at the beginning of the block. Otherwise, copy the 
current context into tc.

void TARGET_SCHED_SET_SCHED_CONTEXT (void *tc) 
[Target Hook] 
Copy target scheduling context pointed to by tc to the current context.

void TARGET_SCHED_CLEAR_SCHED_CONTEXT (void *tc) 
[Target Hook] 
Deallocate internal data in target scheduling context pointed to by tc.

void TARGET_SCHED_FREE_SCHED_CONTEXT (void *tc) 
[Target Hook] 
Deallocate a store for target scheduling context pointed to by tc.

int TARGET_SCHED_SPECULATE_INSN (rtx insn, int request, rtx *new_pat) 
[Target Hook] 
This hook is called by the insn scheduler when insn has only speculative dependencies 
and therefore can be scheduled speculatively. The hook is used to check if the pattern 
of insn has a speculative version and, in case of successful check, to generate that 
speculative pattern. The hook should return 1, if the instruction has a speculative 
form, or −1, if it doesn’t. request describes the type of requested speculation. If the 
return value equals 1 then new_pat is assigned the generated speculative pattern.

bool TARGET_SCHED_NEEDS_BLOCK_P (int dep_status) 
[Target Hook] 
This hook is called by the insn scheduler during generation of recovery code for insn. It 
should return true, if the corresponding check instruction should branch to recovery 
code, or false otherwise.

rtx TARGET_SCHED_GEN_SPEC_CHECK (rtx insn, rtx label, int mutate_p) 
[Target Hook] 
This hook is called by the insn scheduler to generate a pattern for recovery check 
instruction. If mutate_p is zero, then insn is a speculative instruction for which the 
check should be generated. label is either a label of a basic block, where recovery 
code should be emitted, or a null pointer, when requested check doesn’t branch to 
recovery code (a simple check). If mutate_p is nonzero, then a pattern for a branchy 
check corresponding to a simple check denoted by insn should be generated. In this 
case label can’t be null.

bool TARGET_SCHED_FIRST_CYCLE_MULTIPASS_DFA_LOOKAHEAD_GUARD_SPEC 
(const rtx insn) 
[Target Hook] 
This hook is used as a workaround for ‘TARGET_SCHED_FIRST_CYCLE_MULTIPASS_DFA_LOOKAHEAD_GUARD_SPEC’ 
not being called on the first instruction of the ready list. The hook is used to discard 
speculative instructions that stand first in the ready list from being scheduled on the 
current cycle. If the hook returns false, insn will not be chosen to be issued. For
non-speculative instructions, the hook should always return true. For example, in
the ia64 backend the hook is used to cancel data speculative insns when the ALAT

table is nearly full.

```c
void TARGET_SCHED_SET_SCHED_FLAGS (struct spec_info_def *spec_info) {

    This hook is used by the insn scheduler to find out what features should be en-
abled/used. The structure *spec_info should be filled in by the target. The structure
describes speculation types that can be used in the scheduler.
}
```

```c
int TARGET_SCHED_SMS_RES_MII (struct ddg *g) {

    This hook is called by the swing modulo scheduler to calculate a resource-based lower
bound which is based on the resources available in the machine and the resources
required by each instruction. The target backend can use g to calculate such bound.
A very simple lower bound will be used in case this hook is not implemented: the
total number of instructions divided by the issue rate.
}
```

```c
bool TARGET_SCHED_DISPATCH (rtx insn, int x) {

    This hook is called by Haifa Scheduler. It returns true if dispatch scheduling is
supported in hardware and the condition specified in the parameter is true.
}
```

```c
void TARGET_SCHED_DISPATCH_DO (rtx insn, int x) {

    This hook is called by Haifa Scheduler. It performs the operation specified in its
second parameter.
}
```

### 17.19 Dividing the Output into Sections (Texts, Data, . . .)

An object file is divided into sections containing different types of data. In the most common
case, there are three sections: the text section, which holds instructions and read-only data;
the data section, which holds initialized writable data; and the bss section, which holds
uninitialized data. Some systems have other kinds of sections.

‘varasm.c’ provides several well-known sections, such as text_section, data_section
and bss_section. The normal way of controlling a foo_section variable is to define the
associated FOO_SECTION_ASM_OP macro, as described below. The macros are only read
once, when ‘varasm.c’ initializes itself, so their values must be run-time constants. They
may however depend on command-line flags.

*Note:* Some run-time files, such ‘crtstuff.c’, also make use of the FOO_SECTION_ASM_OP
macros, and expect them to be string literals.

Some assemblers require a different string to be written every time a section is selected. If
your assembler falls into this category, you should define the TARGET_ASM_INIT_SECTIONS
hook and use get_unnamed_section to set up the sections.

You must always create a text_section, either by defining TEXT_SECTION_ASM_OP or
by initializing text_section in TARGET_ASM_INIT_SECTIONS. The same is true of data_
section and DATA_SECTION_ASM_OP. If you do not create a distinct readonly_data_
section, the default is to reuse text_section.

All the other ‘varasm.c’ sections are optional, and are null if the target does not provide
them.
TEXT_SECTION_ASM_OP
A C expression whose value is a string, including spacing, containing the assembler operation that should precede instructions and read-only data. Normally "\t.text" is right.

HOT_TEXT_SECTION_NAME
If defined, a C string constant for the name of the section containing most frequently executed functions of the program. If not defined, GCC will provide a default definition if the target supports named sections.

UNLIKELY_EXECUTED_TEXT_SECTION_NAME
If defined, a C string constant for the name of the section containing unlikely executed functions in the program.

DATA_SECTION_ASM_OP
A C expression whose value is a string, including spacing, containing the assembler operation to identify the following data as writable initialized data. Normally "\t.data" is right.

SDATA_SECTION_ASM_OP
If defined, a C expression whose value is a string, including spacing, containing the assembler operation to identify the following data as initialized, writable small data.

READONLY_DATA_SECTION_ASM_OP
A C expression whose value is a string, including spacing, containing the assembler operation to identify the following data as read-only initialized data.

BSS_SECTION_ASM_OP
If defined, a C expression whose value is a string, including spacing, containing the assembler operation to identify the following data as uninitialized global data. If not defined, and neither ASM_OUTPUT_BSS nor ASM_OUTPUTAligned_BSS are defined, uninitialized global data will be output in the data section if '-fno-common' is passed, otherwise ASM_OUTPUT_COMMON will be used.

SBSS_SECTION_ASM_OP
If defined, a C expression whose value is a string, including spacing, containing the assembler operation to identify the following data as uninitialized, writable small data.

TLS_COMMON_ASM_OP
If defined, a C expression whose value is a string containing the assembler operation to identify the following data as thread-local common data. The default is ".tls_common".

TLS_SECTION_ASM_FLAG
If defined, a C expression whose value is a character constant containing the flag used to mark a section as a TLS section. The default is 'T'.

INIT_SECTION_ASM_OP
If defined, a C expression whose value is a string, including spacing, containing the assembler operation to identify the following data as initialization code. If not defined,
GCC will assume such a section does not exist. This section has no corresponding `init_section` variable; it is used entirely in runtime code.

**FINI_SECTION_ASM_OP** ([Macro])

If defined, a C expression whose value is a string, including spacing, containing the assembler operation to identify the following data as finalization code. If not defined, GCC will assume such a section does not exist. This section has no corresponding `fini_section` variable; it is used entirely in runtime code.

**INIT_ARRAY_SECTION_ASM_OP** ([Macro])

If defined, a C expression whose value is a string, including spacing, containing the assembler operation to identify the following data as part of the `.init_array` (or equivalent) section. If not defined, GCC will assume such a section does not exist. Do not define both this macro and `INIT_SECTION_ASM_OP`.

**FINI_ARRAY_SECTION_ASM_OP** ([Macro])

If defined, a C expression whose value is a string, including spacing, containing the assembler operation to identify the following data as part of the `.fini_array` (or equivalent) section. If not defined, GCC will assume such a section does not exist. Do not define both this macro and `FINI_SECTION_ASM_OP`.

**CRT_CALL_STATIC_FUNCTION** (*section_op*, *function*) ([Macro])

If defined, an ASM statement that switches to a different section via `section_op`, calls `function`, and switches back to the text section. This is used in `crtstuff.c` if `INIT_SECTION_ASM_OP` or `FINI_SECTION_ASM_OP` to calls to initialization and finalization functions from the init and fini sections. By default, this macro uses a simple function call. Some ports need hand-crafted assembly code to avoid dependencies on registers initialized in the function prologue or to ensure that constant pools don’t end up too far way in the text section.

**TARGET_LIBGCC_SDATA_SECTION** ([Macro])

If defined, a string which names the section into which small variables defined in `crtstuff` and `libgcc` should go. This is useful when the target has options for optimizing access to small data, and you want the `crtstuff` and `libgcc` routines to be conservative in what they expect of your application yet liberal in what your application expects. For example, for targets with a `.sdata` section (like MIPS), you could compile `crtstuff` with `-G 0` so that it doesn’t require small data support from your application, but use this macro to put small data into `.sdata` so that your application can access these variables whether it uses small data or not.

**FORCE_CODE_SECTION_ALIGN** ([Macro])

If defined, an ASM statement that aligns a code section to some arbitrary boundary. This is used to force all fragments of the `.init` and `.fini` sections to have the same alignment and thus prevent the linker from having to add any padding.

**JUMP_TABLES_IN_TEXT_SECTION** ([Macro])

Define this macro to be an expression with a nonzero value if jump tables (for `tablejump` insns) should be output in the text section, along with the assembler instructions. Otherwise, the readonly data section is used. This macro is irrelevant if there is no separate readonly data section.
void TARGET_ASM_INIT_SECTIONS (void)  
[Target Hook]
Define this hook if you need to do something special to set up the `varasm.c` sections, or if your target has some special sections of its own that you need to create.

GCC calls this hook after processing the command line, but before writing any assembly code, and before calling any of the section-returning hooks described below.

int TARGET_ASM_RELOC_RW_MASK (void)  
[Target Hook]
Return a mask describing how relocations should be treated when selecting sections. Bit 1 should be set if global relocations should be placed in a read-write section; bit 0 should be set if local relocations should be placed in a read-write section.

The default version of this function returns 3 when `-fpic` is in effect, and 0 otherwise. The hook is typically redefined when the target cannot support (some kinds of) dynamic relocations in read-only sections even in executables.

section * TARGET_ASM_SELECT_SECTION (tree exp, int reloc,  
unsigned HOST_WIDE_INT align)  
[Target Hook]
Return the section into which `exp` should be placed. You can assume that `exp` is either a VAR_DECL node or a constant of some sort. `reloc` indicates whether the initial value of `exp` requires link-time relocations. Bit 0 is set when variable contains local relocations only, while bit 1 is set for global relocations. `align` is the constant alignment in bits.

The default version of this function takes care of putting read-only variables in readonly_data_section.

See also USE_SELECT_SECTION_FOR_FUNCTIONS.

USE_SELECT_SECTION_FOR_FUNCTIONS  
[Macro]
Define this macro if you wish TARGET_ASM_SELECT_SECTION to be called for FUNCTION_DECLS as well as for variables and constants.

In the case of a FUNCTION_DECL, `reloc` will be zero if the function has been determined to be likely to be called, and nonzero if it is unlikely to be called.

void TARGET_ASM_UNIQUE_SECTION (tree decl, int reloc)  
[Target Hook]
Build up a unique section name, expressed as a STRING_CST node, and assign it to `DECL_SECTION_NAME (decl)`.

The default version of this function appends the symbol name to the ELF section name that would normally be used for the symbol. For example, the function `foo` would be placed in `.text.foo`. Whatever the actual target object format, this is often good enough.

section * TARGET_ASM_FUNCTION_RODATA_SECTION (tree decl)  
[Target Hook]
Return the readonly data section associated with `DECL_SECTION_NAME (decl)`.

The default version of this function selects `.gnu.linkonce.r.name` if the function’s section is `.gnu.linkonce.t.name`, `.rodata.name` if function is in `.text.name`, and the normal readonly-data section otherwise.
section * TARGET_ASM_SELECT_RTX_SECTION (enum machine_mode mode, rtx x, unsigned HOST_WIDE_INT align)
Return the section into which a constant x, of mode mode, should be placed. You can assume that x is some kind of constant in RTL. The argument mode is redundant except in the case of a const_int rtx. align is the constant alignment in bits.
The default version of this function takes care of putting symbolic constants in flag_pic mode in data_section and everything else in readonly_data_section.

tree TARGET_MANGLE_DECL_ASSEMBLER_NAME (tree decl, tree id)
Define this hook if you need to postprocess the assembler name generated by target-independent code. The id provided to this hook will be the computed name (e.g., the macro DECL_NAME of the decl in C, or the mangled name of the decl in C++). The return value of the hook is an IDENTIFIER_NODE for the appropriate mangled name on your target system. The default implementation of this hook just returns the id provided.

void TARGET_ENCODE_SECTION_INFO (tree decl, rtx rtl, int new_decl_p)
Define this hook if references to a symbol or a constant must be treated differently depending on something about the variable or function named by the symbol (such as what section it is in).
The hook is executed immediately after rtl has been created for decl, which may be a variable or function declaration or an entry in the constant pool. In either case, rtl is the rtl in question. Do not use DECL_RTL (decl) in this hook; that field may not have been initialized yet.

In the case of a constant, it is safe to assume that the rtl is a mem whose address is a symbol_ref. Most decls will also have this form, but that is not guaranteed. Global register variables, for instance, will have a reg for their rtl. (Normally the right thing to do with such unusual rtl is leave it alone.)

The new_decl_p argument will be true if this is the first time that TARGET_ENCODE_SECTION_INFO has been invoked on this decl. It will be false for subsequent invocations, which will happen for duplicate declarations. Whether or not anything must be done for the duplicate declaration depends on whether the hook examines DECL_ATTRIBUTES. new_decl_p is always true when the hook is called for a constant.

The usual thing for this hook to do is to record flags in the symbol_ref, using SYMBOL_REF_FLAG or SYMBOL_REF_FLAGS. Historically, the name string was modified if it was necessary to encode more than one bit of information, but this practice is now discouraged; use SYMBOL_REF_FLAGS.

The default definition of this hook, default_encode_section_info in ‘varasm.c’, sets a number of commonly-useful bits in SYMBOL_REF_FLAGS. Check whether the default does what you need before overriding it.

const char * TARGET_STRIP_NAME_ENCODING (const char *name)
Decode name and return the real name part, sans the characters that TARGET_ENCODE_SECTION_INFO may have added.
bool TARGET_IN_SMALL_DATA_P (const tree exp)  [Target Hook]
Returns true if exp should be placed into a “small data” section. The default version of this hook always returns false.

bool TARGET_HAVE_SRODATA_SECTION  [Target Hook]
Contains the value true if the target places read-only “small data” into a separate section. The default value is false.

bool TARGET_PROFILE_BEFORE_PROLOGUE (void)  [Target Hook]
It returns true if target wants profile code emitted before prologue.
The default version of this hook use the target macro PROFILE_BEFORE_PROLOGUE.

bool TARGET_BINDS_LOCAL_P (const tree exp)  [Target Hook]
Returns true if exp names an object for which name resolution rules must resolve to the current “module” (dynamic shared library or executable image).
The default version of this hook implements the name resolution rules for ELF, which has a looser model of global name binding than other currently supported object file formats.

bool TARGET_HAVE_TLS  [Target Hook]
Contains the value true if the target supports thread-local storage. The default value is false.

17.20 Position Independent Code
This section describes macros that help implement generation of position independent code. Simply defining these macros is not enough to generate valid PIC; you must also add support to the hook TARGET_LEGITIMATE_ADDRESS_P and to the macro PRINT_OPERAND_ADDRESS, as well as LEGITIMIZE_ADDRESS. You must modify the definition of ‘movsi’ to do something appropriate when the source operand contains a symbolic address. You may also need to alter the handling of switch statements so that they use relative addresses.

PIC_OFFSET_TABLE_REGNUM  [Macro]
The register number of the register used to address a table of static data addresses in memory. In some cases this register is defined by a processor’s “application binary interface” (ABI). When this macro is defined, RTL is generated for this register once, as with the stack pointer and frame pointer registers. If this macro is not defined, it is up to the machine-dependent files to allocate such a register (if necessary). Note that this register must be fixed when in use (e.g. when flag_pic is true).

PIC_OFFSET_TABLE_REG_CALL_Clobbered  [Macro]
A C expression that is nonzero if the register defined by PIC_OFFSET_TABLE_REGNUM is clobbered by calls. If not defined, the default is zero. Do not define this macro if PIC_OFFSET_TABLE_REGNUM is not defined.

LEGITIMATE_PIC_OPERAND_P (x)  [Macro]
A C expression that is nonzero if x is a legitimate immediate operand on the target machine when generating position independent code. You can assume that x satisfies CONSTANT_P, so you need not check this. You can also assume flag_pic is true, so you need not check it either. You need not define this macro if all constants (including SYMBOL_REF) can be immediate operands when generating position independent code.
17.21 Defining the Output Assembler Language

This section describes macros whose principal purpose is to describe how to write instructions in assembler language—rather than what the instructions do.

17.21.1 The Overall Framework of an Assembler File

This describes the overall framework of an assembly file.

```c
void TARGET_ASM_FILE_START (void)  
Output to asm_out_file any text which the assembler expects to find at the beginning of a file. The default behavior is controlled by two flags, documented below. Unless your target’s assembler is quite unusual, if you override the default, you should call default_file_start at some point in your target hook. This lets other target files rely on these variables.

bool TARGET_ASM_FILE_START_APP_OFF  
If this flag is true, the text of the macro ASM_APP_OFF will be printed as the very first line in the assembly file, unless ‘-fverbose-asm’ is in effect. (If that macro has been defined to the empty string, this variable has no effect.) With the normal definition of ASM_APP_OFF, the effect is to notify the GNU assembler that it need not bother stripping comments or extra whitespace from its input. This allows it to work a bit faster.

The default is false. You should not set it to true unless you have verified that your port does not generate any extra whitespace or comments that will cause GAS to issue errors in NO_APP mode.

bool TARGET_ASM_FILE_START_FILE_DIRECTIVE  
If this flag is true, output_file_directive will be called for the primary source file, immediately after printing ASM_APP_OFF (if that is enabled). Most ELF assemblers expect this to be done. The default is false.

void TARGET_ASM_FILE_END (void)  
Output to asm_out_file any text which the assembler expects to find at the end of a file. The default is to output nothing.

void file_end_indicate_exec_stack ()  
Some systems use a common convention, the ‘.note.GNU-stack’ special section, to indicate whether or not an object file relies on the stack being executable. If your system uses this convention, you should define TARGET_ASM_FILE_END to this function. If you need to do other things in that hook, have your hook function call this function.

void TARGET_ASM_LTO_START (void)  
Output to asm_out_file any text which the assembler expects to find at the start of an LTO section. The default is to output nothing.

void TARGET_ASM_LTO_END (void)  
Output to asm_out_file any text which the assembler expects to find at the end of an LTO section. The default is to output nothing.
```
void TARGET_ASM_CODE_END (void) [Target Hook]
  Output to asm_out_file any text which is needed before emitting unwind info and debug info at the end of a file. Some targets emit here PIC setup thunks that cannot be emitted at the end of file, because they couldn’t have unwind info then. The default is to output nothing.

ASM_COMMENT_START [Macro]
  A C string constant describing how to begin a comment in the target assembler language. The compiler assumes that the comment will end at the end of the line.

ASM_APP_ON [Macro]
  A C string constant for text to be output before each asm statement or group of consecutive ones. Normally this is "#APP", which is a comment that has no effect on most assemblers but tells the GNU assembler that it must check the lines that follow for all valid assembler constructs.

ASM_APP_OFF [Macro]
  A C string constant for text to be output after each asm statement or group of consecutive ones. Normally this is "#NO_APP", which tells the GNU assembler to resume making the time-saving assumptions that are valid for ordinary compiler output.

ASM_OUTPUT_SOURCE_FILENAME (stream, name) [Macro]
  A C statement to output COFF information or DWARF debugging information which indicates that filename name is the current source file to the stdio stream stream.
  This macro need not be defined if the standard form of output for the file format in use is appropriate.

void TARGET_ASM_OUTPUT_SOURCE_FILENAME (FILE *file, const char *name) [Target Hook]
  Output COFF information or DWARF debugging information which indicates that filename name is the current source file to the stdio stream file.
  This target hook need not be defined if the standard form of output for the file format in use is appropriate.

OUTPUT_QUOTED_STRING (stream, string) [Macro]
  A C statement to output the string string to the stdio stream stream. If you do not call the function output_quoted_string in your config files, GCC will only call it to output filenames to the assembler source. So you can use it to canonicalize the format of the filename using this macro.

ASM_OUTPUT_IDENT (stream, string) [Macro]
  A C statement to output something to the assembler file to handle a ‘#ident’ directive containing the text string. If this macro is not defined, nothing is output for a ‘#ident’ directive.

void TARGET_ASM_NAMED_SECTION (const char *name, unsigned int flags, tree decl) [Target Hook]
  Output assembly directives to switch to section name. The section should have attributes as specified by flags, which is a bit mask of the SECTION_* flags defined in ‘output.h’. If decl is non-NULL, it is the VARDECL or FUNCTION_DECL with which this section is associated.
section * TARGET_ASM_FUNCTION_SECTION (tree decl, enum node, frequency freq, bool startup, bool exit)
    [Target Hook]
    Return preferred text (sub)section for function decl. Main purpose of this function is to separate cold, normal and hot functions. startup is true when function is known to be used only at startup (from static constructors or it is main()). exit is true when function is known to be used only at exit (from static destructors). Return NULL if function should go to default text section.

void TARGET_ASM_FUNCTION_SWITCHED_TEXT_SECTIONS (FILE *file, tree decl, bool new_is_cold)
    [Target Hook]
    Used by the target to emit any assembler directives or additional labels needed when a function is partitioned between different sections. Output should be written to file. The function decl is available as decl and the new section is ‘cold’ if new_is_cold is true.

bool TARGET_HAVE_NAMED_SECTIONS
    [Target Hook]
    This flag is true if the target supports TARGET_ASM_NAMED_SECTION. It must not be modified by command-line option processing.

bool TARGET_HAVE_SWITCHABLE_BSS_SECTIONS
    [Target Hook]
    This flag is true if we can create zeroed data by switching to a BSS section and then using ASM_OUTPUT_SKIP to allocate the space. This is true on most ELF targets.

unsigned int TARGET_SECTION_TYPE_FLAGS (tree decl, const char *name, int reloc)
    [Target Hook]
    Choose a set of section attributes for use by TARGET_ASM_NAMED_SECTION based on a variable or function decl, a section name, and whether or not the declaration’s initializer may contain runtime relocations. decl may be null, in which case read-write data should be assumed.

    The default version of this function handles choosing code vs data, read-only vs read-write data, and flag_pic. You should only need to override this if your target has special flags that might be set via __attribute__.

int TARGET_ASM_RECORD_GCC_SWITCHES (print_switch_type type, const char *text)
    [Target Hook]
    Provides the target with the ability to record the gcc command line switches that have been passed to the compiler, and options that are enabled. The type argument specifies what is being recorded. It can take the following values:

    SWITCH_TYPE_PASSED
        text is a command line switch that has been set by the user.

    SWITCH_TYPE_ENABLED
        text is an option which has been enabled. This might be as a direct result of a command line switch, or because it is enabled by default or because it has been enabled as a side effect of a different command line switch. For example, the ‘-O2’ switch enables various different individual optimization passes.
SWITCH_TYPE_DESCRIPTIVE

text is either NULL or some descriptive text which should be ignored. If text is NULL then it is being used to warn the target hook that either recording is starting or ending. The first time type is SWITCH_TYPE_DESCRIPTIVE and text is NULL, the warning is for start up and the second time the warning is for wind down. This feature is to allow the target hook to make any necessary preparations before it starts to record switches and to perform any necessary tidying up after it has finished recording switches.

SWITCH_TYPE_LINE_START

This option can be ignored by this target hook.

SWITCH_TYPE_LINE_END

This option can be ignored by this target hook.

The hook's return value must be zero. Other return values may be supported in the future.

By default this hook is set to NULL, but an example implementation is provided for ELF based targets. Called elf_record_gcc_switches, it records the switches as ASCII text inside a new, string mergeable section in the assembler output file. The name of the new section is provided by the TARGET_ASM_RECORD_GCC_SWITCHES_SECTION target hook.

```
const char * TARGET_ASM_RECORD_GCC_SWITCHES_SECTION [Target Hook]
```

This is the name of the section that will be created by the example ELF implementation of the TARGET_ASM_RECORD_GCC_SWITCHES target hook.

### 17.21.2 Output of Data

```
const char * TARGET_ASM_BYTE_OP [Target Hook]
const char * TARGET_ASM_ALIGNED_HI_OP [Target Hook]
const char * TARGET_ASM_ALIGNED_SI_OP [Target Hook]
const char * TARGET_ASM_ALIGNED_DI_OP [Target Hook]
const char * TARGET_ASM_ALIGNED_TI_OP [Target Hook]
const char * TARGET_ASM_UNALIGNED_HI_OP [Target Hook]
const char * TARGET_ASM_UNALIGNED_SI_OP [Target Hook]
const char * TARGET_ASM_UNALIGNED_DI_OP [Target Hook]
const char * TARGET_ASM_UNALIGNED_TI_OP [Target Hook]
```

These hooks specify assembly directives for creating certain kinds of integer object. The TARGET_ASM_BYTE_OP directive creates a byte-sized object, the TARGET_ASM_ALIGNED_HI_OP one creates an aligned two-byte object, and so on. Any of the hooks may be NULL, indicating that no suitable directive is available.

The compiler will print these strings at the start of a new line, followed immediately by the object’s initial value. In most cases, the string should contain a tab, a pseudo-op, and then another tab.
bool TARGET_ASM_INTEGER (rtx x, unsigned int size, int aligned_p)  
   [Target Hook]
   The assemble_integer function uses this hook to output an integer object. x is the 
   object’s value, size is its size in bytes and aligned_p indicates whether it is aligned. 
   The function should return true if it was able to output the object. If it returns false, 
   assemble_integer will try to split the object into smaller parts.
   The default implementation of this hook will use the TARGET_ASM_BYTE_OP family of 
   strings, returning false when the relevant string is NULL.

bool TARGET_ASM_OUTPUT_ADDR_CONST_EXTRA (FILE *file, rtx x)  
   [Target Hook]
   A target hook to recognize rtx patterns that output_addr_const can’t deal with, 
   and output assembly code to file corresponding to the pattern x. This may be used 
   to allow machine-dependent UNSPECs to appear within constants.
   If target hook fails to recognize a pattern, it must return false, so that a standard 
   error message is printed. If it prints an error message itself, by calling, for example, 
   output_operand_lossage, it may just return true.

OUTPUT_ADDR_CONST_EXTRA (stream, x, fail)  
   [Macro]
   A C statement to recognize rtx patterns that output_addr_const can’t deal with, 
   and output assembly code to stream corresponding to the pattern x. This may be used 
   to allow machine-dependent UNSPECs to appear within constants.
   If OUTPUT_ADDR_CONST_EXTRA fails to recognize a pattern, it must goto fail, so that 
   a standard error message is printed. If it prints an error message itself, by calling, for 
   example, output_operand_lossage, it may just complete normally.

ASM_OUTPUT_ASCII (stream, ptr, len)  
   [Macro]
   A C statement to output to the stdio stream stream an assembler instruction to 
   assemble a string constant containing the len bytes at ptr. ptr will be a C expression 
   of type char * and len a C expression of type int.
   If the assembler has a .ascii pseudo-op as found in the Berkeley Unix assembler, do 
   not define the macro ASM_OUTPUT_ASCII.

ASM_OUTPUT_FDESC (stream, decl, n)  
   [Macro]
   A C statement to output word n of a function descriptor for decl. This must be 
   defined if TARGET_VTABLE_USES_DESCRIPTORS is defined, and is otherwise unused.

CONSTANT_POOL_BEFORE_FUNCTION  
   [Macro]
   You may define this macro as a C expression. You should define the expression to 
   have a nonzero value if GCC should output the constant pool for a function before 
   the code for the function, or a zero value if GCC should output the constant pool 
   after the function. If you do not define this macro, the usual case, GCC will output 
   the constant pool before the function.

ASM_OUTPUT_POOL_PROLOGUE (file, funname, fundecl, size)  
   [Macro]
   A C statement to output assembler commands to define the start of the constant pool 
   for a function. funname is a string giving the name of the function. Should the return
type of the function be required, it can be obtained via `fundecl`. `size` is the size, in bytes, of the constant pool that will be written immediately after this call. If no constant-pool prefix is required, the usual case, this macro need not be defined.

`ASM_OUTPUT_SPECIAL_POOL_ENTRY (file, x, mode, align, labelno, jump_to)`

A C statement (with or without semicolon) to output a constant in the constant pool, if it needs special treatment. (This macro need not do anything for RTL expressions that can be output normally.)

The argument `file` is the standard I/O stream to output the assembler code on. `x` is the RTL expression for the constant to output, and `mode` is the machine mode (in case `x` is a `const_int`). `align` is the required alignment for the value `x`; you should output an assembler directive to force this much alignment.

The argument `labelno` is a number to use in an internal label for the address of this pool entry. The definition of this macro is responsible for outputting the label definition at the proper place. Here is how to do this:

```c
(*targetm.asm_out.internal_label) (file, "LC", labelno);
```

When you output a pool entry specially, you should end with a `goto` to the label `jump_to`. This will prevent the same pool entry from being output a second time in the usual manner.

You need not define this macro if it would do nothing.

`ASM_OUTPUT_POOL_EPILOGUE (file funname fundecl size)`

A C statement to output assembler commands to at the end of the constant pool for a function. `funname` is a string giving the name of the function. Should the return type of the function be required, you can obtain it via `fundecl`. `size` is the size, in bytes, of the constant pool that GCC wrote immediately before this call.

If no constant-pool epilogue is required, the usual case, you need not define this macro.

`IS_ASM_LOGICAL_LINE_SEPARATOR (C, STR)`

Define this macro as a C expression which is nonzero if `C` is used as a logical line separator by the assembler. `STR` points to the position in the string where `C` was found; this can be used if a line separator uses multiple characters.

If you do not define this macro, the default is that only the character ‘;’ is treated as a logical line separator.

`const char * TARGET_ASM_OPEN_PAREN`  
`const char * TARGET_ASM_CLOSE_PAREN`

These target hooks are C string constants, describing the syntax in the assembler for grouping arithmetic expressions. If not overridden, they default to normal parentheses, which is correct for most assemblers.

These macros are provided by ‘real.h’ for writing the definitions of `ASM_OUTPUT_DOUBLE` and the like:

`REAL_VALUE_TO_TARGET_SINGLE (x, 1)`  
`REAL_VALUE_TO_TARGET_DOUBLE (x, 1)}`
REAL_VALUE_TO_TARGET_LONG_DOUBLE (x, 1) [Macro]
REAL_VALUE_TO_TARGET_DECIMAL32 (x, 1) [Macro]
REAL_VALUE_TO_TARGET_DECIMAL64 (x, 1) [Macro]
REAL_VALUE_TO_TARGET_DECIMAL128 (x, 1) [Macro]

These translate x, of type REAL_VALUE_TYPE, to the target’s floating point representation, and store its bit pattern in the variable l. For REAL_VALUE_TO_TARGET_SINGLE and REAL_VALUE_TO_TARGET_DECIMAL32, this variable should be a simple long int. For the others, it should be an array of long int. The number of elements in this array is determined by the size of the desired target floating point data type: 32 bits of it go in each long int array element. Each array element holds 32 bits of the result, even if long int is wider than 32 bits on the host machine.

The array element values are designed so that you can print them out using fprintf in the order they should appear in the target machine’s memory.

### 17.21.3 Output of Uninitialized Variables

Each of the macros in this section is used to do the whole job of outputting a single uninitialized variable.

ASM_OUTPUT_COMMON (stream, name, size, rounded) [Macro]
A C statement (sans semicolon) to output to the stdio stream stream the assembler definition of a common-label named name whose size is size bytes. The variable rounded is the size rounded up to whatever alignment the caller wants. It is possible that size may be zero, for instance if a struct with no other member than a zero-length array is defined. In this case, the backend must output a symbol definition that allocates at least one byte, both so that the address of the resulting object does not compare equal to any other, and because some object formats cannot even express the concept of a zero-sized common symbol, as that is how they represent an ordinary undefined external.

Use the expression assemble_name (stream, name) to output the name itself; before and after that, output the additional assembler syntax for defining the name, and a newline.

This macro controls how the assembler definitions of uninitialized common global variables are output.

ASM_OUTPUT_ALIGNED_COMMON (stream, name, size, alignment) [Macro]
Like ASM_OUTPUT_COMMON except takes the required alignment as a separate, explicit argument. If you define this macro, it is used in place of ASM_OUTPUT_COMMON, and gives you more flexibility in handling the required alignment of the variable. The alignment is specified as the number of bits.

ASM_OUTPUT_ALIGNED_DECL_COMMON (stream, decl, name, size, alignment) [Macro]
Like ASM_OUTPUT_ALIGNED_COMMON except that decl of the variable to be output, if there is one, or NULL_TREE if there is no corresponding variable. If you define this macro, GCC will use it in place of both ASM_OUTPUT_COMMON and ASM_OUTPUT_ALIGNED_COMMON. Define this macro when you need to see the variable’s decl in order to chose what to output.
ASM_OUTPUT_BSS (stream, decl, name, size, rounded) [Macro]

A C statement (sans semicolon) to output to the stdio stream stream the assembler definition of uninitialized global decl named name whose size is size bytes. The variable rounded is the size rounded up to whatever alignment the caller wants.

Try to use function asm_output_bss defined in ‘varasm.c’ when defining this macro. If unable, use the expression assemble_name (stream, name) to output the name itself; before and after that, output the additional assembler syntax for defining the name, and a newline.

There are two ways of handling global BSS. One is to define either this macro or its aligned counterpart, ASM_OUTPUT_ALIGNED_BSS. The other is to have TARGET_ASM_SELECT_SECTION return a switchable BSS section (see [TARGET_HAVE_SWITCHABLE_BSS_SECTIONS], page 499). You do not need to do both.

Some languages do not have common data, and require a non-common form of global BSS in order to handle uninitialized globals efficiently. C++ is one example of this. However, if the target does not support global BSS, the front end may choose to make globals common in order to save space in the object file.

ASM_OUTPUT_ALIGNED_BSS (stream, decl, name, size, alignment) [Macro]

Like ASM_OUTPUT_BSS except takes the required alignment as a separate, explicit argument. If you define this macro, it is used in place of ASM_OUTPUT_BSS, and gives you more flexibility in handling the required alignment of the variable. The alignment is specified as the number of bits.

Try to use function asm_output_aligned_bss defined in file ‘varasm.c’ when defining this macro.

ASM_OUTPUT_LOCAL (stream, name, size, rounded) [Macro]

A C statement (sans semicolon) to output to the stdio stream stream the assembler definition of a local-common-label named name whose size is size bytes. The variable rounded is the size rounded up to whatever alignment the caller wants.

Use the expression assemble_name (stream, name) to output the name itself; before and after that, output the additional assembler syntax for defining the name, and a newline.

This macro controls how the assembler definitions of uninitialized static variables are output.

ASM_OUTPUTAligned_LOCAL (stream, name, size, alignment) [Macro]

Like ASM_OUTPUT_LOCAL except takes the required alignment as a separate, explicit argument. If you define this macro, it is used in place of ASM_OUTPUT_LOCAL, and gives you more flexibility in handling the required alignment of the variable. The alignment is specified as the number of bits.

ASM_OUTPUTAligned_DECL_LOCAL (stream, decl, name, size, alignment) [Macro]

Like ASM_OUTPUTAligned_DECL except that decl of the variable to be output, if there is one, or NULL_TREE if there is no corresponding variable. If you define this macro, GCC will use it in place of both ASM_OUTPUT_DECL and ASM_OUTPUTAlignedDECL.
Define this macro when you need to see the variable’s decl in order to chose what to output.

### 17.21.4 Output and Generation of Labels

This is about outputting labels.

**ASM_OUTPUT_LABEL (stream, name)**  
A C statement (sans semicolon) to output to the stdio stream `stream` the assembler definition of a label named `name`. Use the expression `assemble_name (stream, name)` to output the name itself; before and after that, output the additional assembler syntax for defining the name, and a newline. A default definition of this macro is provided which is correct for most systems.

**ASM_OUTPUT_FUNCTION_LABEL (stream, name, decl)**  
A C statement (sans semicolon) to output to the stdio stream `stream` the assembler definition of a label named `name` of a function. Use the expression `assemble_name (stream, name)` to output the name itself; before and after that, output the additional assembler syntax for defining the name, and a newline. A default definition of this macro is provided which is correct for most systems.

If this macro is not defined, then the function name is defined in the usual manner as a label (by means of `ASM_OUTPUT_LABEL`).

**ASM_OUTPUT_INTERNAL_LABEL (stream, name)**  
Identical to `ASM_OUTPUT_LABEL`, except that `name` is known to refer to a compiler-generated label. The default definition uses `assemble_name_raw`, which is like `assemble_name` except that it is more efficient.

**SIZE_ASM_OP**  
A C string containing the appropriate assembler directive to specify the size of a symbol, without any arguments. On systems that use ELF, the default (in `config/elfos.h`) is `"\t.size\t"`; on other systems, the default is not to define this macro.

Define this macro only if it is correct to use the default definitions of `ASM_OUTPUT_SIZE_DIRECTIVE` and `ASM_OUTPUT_MEASURED_SIZE` for your system. If you need your own custom definitions of those macros, or if you do not need explicit symbol sizes at all, do not define this macro.

**ASM_OUTPUT_SIZE_DIRECTIVE (stream, name, size)**  
A C statement (sans semicolon) to output to the stdio stream `stream` a directive telling the assembler that the size of the symbol `name` is `size`. `size` is a `HOST_WIDE_INT`. If you define `SIZE_ASM_OP`, a default definition of this macro is provided.

**ASM_OUTPUT_MEASURED_SIZE (stream, name)**  
A C statement (sans semicolon) to output to the stdio stream `stream` a directive telling the assembler to calculate the size of the symbol `name` by subtracting its address from the current address.

If you define `SIZE_ASM_OP`, a default definition of this macro is provided. The default assumes that the assembler recognizes a special `\t` symbol as referring to the current
address, and can calculate the difference between this and another symbol. If your assembler does not recognize ‘.’ or cannot do calculations with it, you will need to redefine `ASM_OUTPUT_MEASURED_SIZE` to use some other technique.

`TYPE_ASM_OP`  
[Macro]  
A C string containing the appropriate assembler directive to specify the type of a symbol, without any arguments. On systems that use ELF, the default (in `config/elfos.h`) is “\t.type\t”; on other systems, the default is not to define this macro.

Define this macro only if it is correct to use the default definition of `ASM_OUTPUT_TYPE_DIRECTIVE` for your system. If you need your own custom definition of this macro, or if you do not need explicit symbol types at all, do not define this macro.

`TYPE_OPERAND_FMT`  
[Macro]  
A C string which specifies (using `printf` syntax) the format of the second operand to `TYPE_ASM_OP`. On systems that use ELF, the default (in `config/elfos.h`) is “@%s” ; on other systems, the default is not to define this macro.

Define this macro only if it is correct to use the default definition of `ASM_OUTPUT_TYPE_DIRECTIVE` for your system. If you need your own custom definition of this macro, or if you do not need explicit symbol types at all, do not define this macro.

`ASM_OUTPUT_TYPE_DIRECTIVE (stream, type)`  
[Macro]  
A C statement (sans semicolon) to output to the stdio stream `stream` a directive telling the assembler that the type of the symbol `name` is `type`. `type` is a C string; currently, that string is always either “`function`” or “`object`”, but you should not count on this.

If you define `TYPE_ASM_OP` and `TYPE_OPERAND_FMT`, a default definition of this macro is provided.

`ASM_DECLARE_FUNCTION_NAME (stream, name, decl)`  
[Macro]  
A C statement (sans semicolon) to output to the stdio stream `stream` any text necessary for declaring the name `name` of a function which is being defined. This macro is responsible for outputting the label definition (perhaps using `ASM_OUTPUT_FUNCTION_LABEL`). The argument `decl` is the `FUNCTION_DECL` tree node representing the function.

If this macro is not defined, then the function name is defined in the usual manner as a label (by means of `ASM_OUTPUT_FUNCTION_LABEL`).

You may wish to use `ASM_OUTPUT_TYPE_DIRECTIVE` in the definition of this macro.

`ASM_DECLARE_FUNCTION_SIZE (stream, name, decl)`  
[Macro]  
A C statement (sans semicolon) to output to the stdio stream `stream` any text necessary for declaring the size of a function which is being defined. The argument `name` is the name of the function. The argument `decl` is the `FUNCTION_DECL` tree node representing the function.

If this macro is not defined, then the function size is not defined.

You may wish to use `ASM_OUTPUT_MEASURED_SIZE` in the definition of this macro.
ASM_DECLARE_OBJECT_NAME (stream, name, decl)  [Macro]
A C statement (sans semicolon) to output to the stdio stream stream any text necessary for declaring the name name of an initialized variable which is being defined. This macro must output the label definition (perhaps using ASM_OUTPUT_LABEL). The argument decl is the VARDECL tree node representing the variable.
If this macro is not defined, then the variable name is defined in the usual manner as a label (by means of ASM_OUTPUT_LABEL).
You may wish to use ASM_OUTPUT_TYPE_DIRECTIVE and/or ASM_OUTPUT_SIZE_DIRECTIVE in the definition of this macro.

void TARGET_ASM_DECLARE_CONSTANT_NAME (FILE *file, const char *name, const tree expr, HOST_WIDE_INT size)  [Target Hook]
A target hook to output to the stdio stream file any text necessary for declaring the name name of a constant which is being defined. This target hook is responsible for outputting the label definition (perhaps using assemble_label). The argument expr is the value of the constant, and size is the size of the constant in bytes. The name will be an internal label.
The default version of this target hook, define the name in the usual manner as a label (by means of assemble_label).
You may wish to use ASM_OUTPUT_TYPE_DIRECTIVE in this target hook.

ASM_DECLARE_REGISTER_GLOBAL (stream, decl, regno, name)  [Macro]
A C statement (sans semicolon) to output to the stdio stream stream any text necessary for claiming a register regno for a global variable decl with name name.
If you don’t define this macro, that is equivalent to defining it to do nothing.

ASM_FINISH_DECLARE_OBJECT (stream, decl, toplevel, atend)  [Macro]
A C statement (sans semicolon) to finish up declaring a variable name once the compiler has processed its initializer fully and thus has had a chance to determine the size of an array when controlled by an initializer. This is used on systems where it’s necessary to declare something about the size of the object.
If you don’t define this macro, that is equivalent to defining it to do nothing.
You may wish to use ASM_OUTPUT_SIZE_DIRECTIVE and/or ASM_OUTPUT_MEASURED_SIZE in the definition of this macro.

void TARGET_ASM_GLOBALIZE_LABEL (FILE *stream, const char *name)  [Target Hook]
This target hook is a function to output to the stdio stream stream some commands that will make the label name global; that is, available for reference from other files.
The default implementation relies on a proper definition of GLOBAL_ASM_OP.

void TARGET_ASM_GLOBALIZE_DECL_NAME (FILE *stream, tree decl)  [Target Hook]
This target hook is a function to output to the stdio stream stream some commands that will make the name associated with decl global; that is, available for reference from other files.
The default implementation uses the TARGET_ASM_GLOBALIZE_LABEL target hook.
**ASM_WEAKEN_LABEL** *(stream, name)*  
A C statement (sans semicolon) to output to the stdio stream `stream` some commands that will make the label `name` weak; that is, available for reference from other files but only used if no other definition is available. Use the expression `assemble_name(stream, name)` to output the name itself; before and after that, output the additional assembler syntax for making that name weak, and a newline.

If you don’t define this macro or **ASM_WEAKEN_DECL**, GCC will not support weak symbols and you should not define the **SUPPORTS_WEAK** macro.

**ASM_WEAKEN_DECL** *(stream, decl, name, value)*  
Combines (and replaces) the function of **ASM_WEAKEN_LABEL** and **ASM_OUTPUT_WEAK_ALIAS**, allowing access to the associated function or variable decl. If `value` is not `NULL`, this C statement should output to the stdio stream `stream` assembler code which defines (equates) the weak symbol `name` to have the value `value`. If `value` is `NULL`, it should output commands to make `name` weak.

**ASM_OUTPUT_WEAKREF** *(stream, decl, name, value)*  
Outputs a directive that enables `name` to be used to refer to symbol `value` with weak-symbol semantics. `decl` is the declaration of `name`.

**SUPPORTS_WEAK**  
A preprocessor constant expression which evaluates to true if the target supports weak symbols.

If you don’t define this macro, ‘defaults.h’ provides a default definition. If either **ASM_WEAKEN_LABEL** or **ASM_WEAKEN_DECL** is defined, the default definition is ‘1’; otherwise, it is ‘0’.

**TARGET_SUPPORTS_WEAK**  
A C expression which evaluates to true if the target supports weak symbols.

If you don’t define this macro, ‘defaults.h’ provides a default definition. The default definition is ‘(SUPPORTS_WEAK)’. Define this macro if you want to control weak symbol support with a compiler flag such as ‘-melf’.

**MAKE_DECL_ONE_ONLY** *(decl)*  
A C statement (sans semicolon) to mark `decl` to be emitted as a public symbol such that extra copies in multiple translation units will be discarded by the linker. Define this macro if your object file format provides support for this concept, such as the ‘COMDAT’ section flags in the Microsoft Windows PE/COFF format, and this support requires changes to `decl`, such as putting it in a separate section.

**SUPPORTS_ONE_ONLY**  
A C expression which evaluates to true if the target supports one-only semantics.

If you don’t define this macro, ‘varasm.c’ provides a default definition. If **MAKE_DECL_ONE_ONLY** is defined, the default definition is ‘1’; otherwise, it is ‘0’. Define this macro if you want to control one-only symbol support with a compiler flag, or if setting the **DECL_ONE_ONLY** flag is enough to mark a declaration to be emitted as one-only.
void TARGET_ASM_ASSEMBLE_VISIBILITY (tree decl, int visibility)  
[Target Hook]  
This target hook is a function to output to asm_out_file some commands that will make the symbol(s) associated with decl have hidden, protected or internal visibility as specified by visibility.

TARGET_WEAK_NOT_IN_ARCHIVE_TOC  
[Macro]  
A C expression that evaluates to true if the target’s linker expects that weak symbols do not appear in a static archive’s table of contents. The default is 0.

Leaving weak symbols out of an archive’s table of contents means that, if a symbol will only have a definition in one translation unit and will have undefined references from other translation units, that symbol should not be weak. Defining this macro to be nonzero will thus have the effect that certain symbols that would normally be weak (explicit template instantiations, and vtables for polymorphic classes with noninline key methods) will instead be nonweak.

The C++ ABI requires this macro to be zero. Define this macro for targets where full C++ ABI compliance is impossible and where linker restrictions require weak symbols to be left out of a static archive’s table of contents.

ASM_OUTPUT_EXTERNAL (stream, decl, name)  
[Macro]  
A C statement (sans semicolon) to output to thestdio stream stream any text necessary for declaring the name of an external symbol named name which is referenced in this compilation but not defined. The value of decl is the tree node for the declaration.

This macro need not be defined if it does not need to output anything. The GNU assembler and most Unix assemblers don’t require anything.

void TARGET_ASM_EXTERNAL_LIBCALL (rtx symref)  
[Target Hook]  
This target hook is a function to output to asm_out_file an assembler pseudo-op to declare a library function name external. The name of the library function is given by symref, which is a symbol_ref.

void TARGET_ASM_MARK_DECL_PRESERVED (const char *symbol)  
[Target Hook]  
This target hook is a function to output to asm_out_file an assembler directive to annotate symbol as used. The Darwin target uses the .no_dead_code_strip directive.

ASM_OUTPUT_LABELREF (stream, name)  
[Macro]  
A C statement (sans semicolon) to output to thestdio stream stream a reference in assembler syntax to a label named name. This should add ‘_’ to the front of the name, if that is customary on your operating system, as it is in most Berkeley Unix systems. This macro is used in assemble_name.

tree TARGET_MANGLE_ASSEMBLER_NAME (const char *name)  
[Target Hook]  
Given a symbol name, perform same mangling as varasm.c’s assemble_name, but in memory rather than to a file stream, returning result as an IDENTIFIER_NODE. Required for correct LTO symtabs. The default implementation calls the TARGET_STRIP_NAME_ENCODING hook and then prepends the USER_LABEL_PREFIX, if any.
ASM_OUTPUT_SYMBOL_REF (stream, sym)  [Macro]
A C statement (sans semicolon) to output a reference to SYMBOL_REF sym. If not defined, assemble_name will be used to output the name of the symbol. This macro may be used to modify the way a symbol is referenced depending on information encoded by TARGET_ENCODE_SECTION_INFO.

ASM_OUTPUT_LABEL_REF (stream, buf)  [Macro]
A C statement (sans semicolon) to output a reference to buf, the result of ASM_GENERATE_INTERNAL_LABEL. If not defined, assemble_name will be used to output the name of the symbol. This macro is not used by output_asm_label, or the %l specifier that calls it; the intention is that this macro should be set when it is necessary to output a label differently when its address is being taken.

void TARGET_ASM_INTERNAL_LABEL (FILE *stream, const char *prefix, unsigned long labelno)  [Target Hook]
A function to output to the stdio stream stream a label whose name is made from the string prefix and the number labelno.

It is absolutely essential that these labels be distinct from the labels used for user-level functions and variables. Otherwise, certain programs will have name conflicts with internal labels.

It is desirable to exclude internal labels from the symbol table of the object file. Most assemblers have a naming convention for labels that should be excluded; on many systems, the letter 'L' at the beginning of a label has this effect. You should find out what convention your system uses, and follow it.

The default version of this function utilizes ASM_GENERATE_INTERNAL_LABEL.

ASM_OUTPUT_DEBUG_LABEL (stream, prefix, num)  [Macro]
A C statement to output to the stdio stream stream a debug info label whose name is made from the string prefix and the number num. This is useful for VLIW targets, where debug info labels may need to be treated differently than branch target labels. On some systems, branch target labels must be at the beginning of instruction bundles, but debug info labels can occur in the middle of instruction bundles.

If this macro is not defined, then (*targetm.asm_out.internal_label) will be used.

ASM_GENERATE_INTERNAL_LABEL (string, prefix, num)  [Macro]
A C statement to store into the string string a label whose name is made from the string prefix and the number num.

This string, when output subsequently by assemble_name, should produce the output that (*targetm.asm_out.internal_label) would produce with the same prefix and num.

If the string begins with '*', then assemble_name will output the rest of the string unchanged. It is often convenient for ASM_GENERATE_INTERNAL_LABEL to use '*' in this way. If the string doesn't start with '*', then ASM_OUTPUT_LABELREF gets to output the string, and may change it. (Of course, ASM_OUTPUT_LABELREF is also part of your machine description, so you should know what it does on your machine.)
**ASM_FORMAT_PRIVATE_NAME** *(outvar, name, number)*  
A C expression to assign to `outvar` (which is a variable of type `char *`) a newly allocated string made from the string `name` and the number `number`, with some suitable punctuation added. Use `alloca` to get space for the string.  
The string will be used as an argument to `ASM_OUTPUT_LABELREF` to produce an assembler label for an internal static variable whose name is `name`. Therefore, the string must be such as to result in valid assembler code. The argument `number` is different each time this macro is executed; it prevents conflicts between similarly-named internal static variables in different scopes.  
Ideally this string should not be a valid C identifier, to prevent any conflict with the user’s own symbols. Most assemblers allow periods or percent signs in assembler symbols; putting at least one of these between the name and the number will suffice.  
If this macro is not defined, a default definition will be provided which is correct for most systems.

**ASM_OUTPUT_DEF** *(stream, name, value)*  
A C statement to output to the stdio stream `stream` assembler code which defines (equates) the symbol `name` to have the value `value`.  
If `SET_ASM_OP` is defined, a default definition is provided which is correct for most systems.

**ASM_OUTPUT_DEF_FROM_DECLS** *(stream, decl_of_name, decl_of_value)*  
A C statement to output to the stdio stream `stream` assembler code which defines (equates) the symbol whose tree node is `decl_of_name` to have the value of the tree node `decl_of_value`. This macro will be used in preference to ‘`ASM_OUTPUT_DEF`’ if it is defined and if the tree nodes are available.  
If `SET_ASM_OP` is defined, a default definition is provided which is correct for most systems.

**TARGET_DEFERRED_OUTPUT_DEFS** *(decl_of_name, decl_of_value)*  
A C statement that evaluates to true if the assembler code which defines (equates) the symbol whose tree node is `decl_of_name` to have the value of the tree node `decl_of_value` should be emitted near the end of the current compilation unit. The default is to not defer output of defines. This macro affects defines output by ‘`ASM_OUTPUT_DEF`’ and ‘`ASM_OUTPUT_DEF_FROM_DECLS`’.

**ASM_OUTPUT_WEAK_ALIAS** *(stream, name, value)*  
A C statement to output to the stdio stream `stream` assembler code which defines (equates) the weak symbol `name` to have the value `value`. If `value` is `NULL`, it defines `name` as an undefined weak symbol.  
Define this macro if the target only supports weak aliases; define `ASM_OUTPUT_DEF` instead if possible.

**OBJC_GEN_METHOD_LABEL** *(buf, is_inst, class_name, cat_name, sel_name)*  
Define this macro to override the default assembler names used for Objective-C methods.
The default name is a unique method number followed by the name of the class (e.g. ‘_1_Foo’). For methods in categories, the name of the category is also included in the assembler name (e.g. ‘_1_Foo_Bar’).

These names are safe on most systems, but make debugging difficult since the method’s selector is not present in the name. Therefore, particular systems define other ways of computing names.

buf is an expression of type char * which gives you a buffer in which to store the name; its length is as long as class_name, cat_name and sel_name put together, plus 50 characters extra.

The argument is_inst specifies whether the method is an instance method or a class method; class_name is the name of the class; cat_name is the name of the category (or NULL if the method is not in a category); and sel_name is the name of the selector.

On systems where the assembler can handle quoted names, you can use this macro to provide more human-readable names.

\[
\text{ASM_DECLARE_CLASS_REFERENCE (stream, name)} \quad \text{[Macro]}
\]

A C statement (sans semicolon) to output to the stdio stream stream commands to declare that the label name is an Objective-C class reference. This is only needed for targets whose linkers have special support for NeXT-style runtimes.

\[
\text{ASM_DECLARE_UNRESOLVED_REFERENCE (stream, name)} \quad \text{[Macro]}
\]

A C statement (sans semicolon) to output to the stdio stream stream commands to declare that the label name is an unresolved Objective-C class reference. This is only needed for targets whose linkers have special support for NeXT-style runtimes.

17.21.5 How Initialization Functions Are Handled

The compiled code for certain languages includes constructors (also called initialization routines)—functions to initialize data in the program when the program is started. These functions need to be called before the program is “started”—that is to say, before main is called.

Compiling some languages generates destructors (also called termination routines) that should be called when the program terminates.

To make the initialization and termination functions work, the compiler must output something in the assembler code to cause those functions to be called at the appropriate time. When you port the compiler to a new system, you need to specify how to do this.

There are two major ways that GCC currently supports the execution of initialization and termination functions. Each way has two variants. Much of the structure is common to all four variations.

The linker must build two lists of these functions—a list of initialization functions, called ___CTOR_LIST___, and a list of termination functions, called ___DTOR_LIST___.

Each list always begins with an ignored function pointer (which may hold 0, −1, or a count of the function pointers after it, depending on the environment). This is followed by a series of zero or more function pointers to constructors (or destructors), followed by a function pointer containing zero.
Depending on the operating system and its executable file format, either ‘crtstuff.c’ or ‘libgcc2.c’ traverses these lists at startup time and exit time. Constructors are called in reverse order of the list; destructors in forward order.

The best way to handle static constructors works only for object file formats which provide arbitrarily-named sections. A section is set aside for a list of constructors, and another for a list of destructors. Traditionally these are called ‘.ctors’ and ‘.dtors’. Each object file that defines an initialization function also puts a word in the constructor section to point to that function. The linker accumulates all these words into one contiguous ‘.ctors’ section. Termination functions are handled similarly.

This method will be chosen as the default by ‘target-def.h’ if TARGET_ASM_NAMED_SECTION is defined. A target that does not support arbitrary sections, but does support special designated constructor and destructor sections may define CTORS_SECTION_ASM_OP and DTORS_SECTION_ASM_OP to achieve the same effect.

When arbitrary sections are available, there are two variants, depending upon how the code in ‘crtstuff.c’ is called. On systems that support a .init section which is executed at program startup, parts of ‘crtstuff.c’ are compiled into that section. The program is linked by the gcc driver like this:

```
ld -o output_file crti.o crtbegin.o ... -lgcc crtend.o crtn.o
```

The prologue of a function (__init) appears in the .init section of ‘crti.o’; the epilogue appears in ‘crtn.o’. Likewise for the function __fini in the .fini section. Normally these files are provided by the operating system or by the GNU C library, but are provided by GCC for a few targets.

The objects ‘crtbegin.o’ and ‘crtend.o’ are (for most targets) compiled from ‘crtstuff.c’. They contain, among other things, code fragments within the .init and .fini sections that branch to routines in the .text section. The linker will pull all parts of a section together, which results in a complete __init function that invokes the routines we need at startup.

To use this variant, you must define the INIT_SECTION_ASM_OP macro properly.

If no init section is available, when GCC compiles any function called main (or more accurately, any function designated as a program entry point by the language front end calling expand_main_function), it inserts a procedure call to __main as the first executable code after the function prologue. The __main function is defined in ‘libgcc2.c’ and runs the global constructors.

In file formats that don’t support arbitrary sections, there are again two variants. In the simplest variant, the GNU linker (GNU ld) and an ‘a.out’ format must be used. In this case, TARGET_ASM_CONSTRUCTOR is defined to produce a .stabs entry of type ‘N_SETT’, referencing the name __CTOR_LIST__, and with the address of the void function containing the initialization code as its value. The GNU linker recognizes this as a request to add the value to a set; the values are accumulated, and are eventually placed in the executable as a vector in the format described above, with a leading (ignored) count and a trailing zero element. TARGET_ASM_DESTRUCTOR is handled similarly. Since no init section is available, the absence of INIT_SECTION_ASM_OP causes the compilation of main to call __main as above, starting the initialization process.

The last variant uses neither arbitrary sections nor the GNU linker. This is preferable when you want to do dynamic linking and when using file formats which the GNU linker
does not support, such as ‘ECOFF’. In this case, \texttt{TARGET_HAVE\_CTORS\_DTORS} is false, initialization and termination functions are recognized simply by their names. This requires an extra program in the linkage step, called \texttt{collect2}. This program pretends to be the linker, for use with GCC; it does its job by running the ordinary linker, but also arranges to include the vectors of initialization and termination functions. These functions are called via \texttt{__main} as described above. In order to use this method, \texttt{use\_collect2} must be defined in the target in ‘\texttt{config.gcc}’.

### 17.21.6 Macros Controlling Initialization Routines

Here are the macros that control how the compiler handles initialization and termination functions:

\textbf{\texttt{INIT\_SECTION\_ASM\_OP}} \quad \textbf{[Macro]}

If defined, a C string constant, including spacing, for the assembler operation to identify the following data as initialization code. If not defined, GCC will assume such a section does not exist. When you are using special sections for initialization and termination functions, this macro also controls how ‘\texttt{crtstuff.c}’ and ‘\texttt{libgcc2.c}’ arrange to run the initialization functions.

\textbf{\texttt{HAS\_INIT\_SECTION}} \quad \textbf{[Macro]}

If defined, \texttt{main} will not call \texttt{__main} as described above. This macro should be defined for systems that control start-up code on a symbol-by-symbol basis, such as OSF/1, and should not be defined explicitly for systems that support \texttt{INIT\_SECTION\_ASM\_OP}.

\textbf{\texttt{LD\_INIT\_SWITCH}} \quad \textbf{[Macro]}

If defined, a C string constant for a switch that tells the linker that the following symbol is an initialization routine.

\textbf{\texttt{LD\_FINI\_SWITCH}} \quad \textbf{[Macro]}

If defined, a C string constant for a switch that tells the linker that the following symbol is a finalization routine.

\textbf{\texttt{COLLECT\_SHARED\_INIT\_FUNC (stream, func)}} \quad \textbf{[Macro]}

If defined, a C statement that will write a function that can be automatically called when a shared library is loaded. The function should call \texttt{func}, which takes no arguments. If not defined, and the object format requires an explicit initialization function, then a function called \texttt{GLOBAL\_DI} will be generated.

This function and the following one are used by \texttt{collect2} when linking a shared library that needs constructors or destructors, or has DWARF2 exception tables embedded in the code.

\textbf{\texttt{COLLECT\_SHARED\_FINI\_FUNC (stream, func)}} \quad \textbf{[Macro]}

If defined, a C statement that will write a function that can be automatically called when a shared library is unloaded. The function should call \texttt{func}, which takes no arguments. If not defined, and the object format requires an explicit finalization function, then a function called \texttt{GLOBAL\_DD} will be generated.
INVOKE__main

If defined, main will call __main despite the presence of INIT_SECTION_ASM_OP. This macro should be defined for systems where the init section is not actually run automatically, but is still useful for collecting the lists of constructors and destructors.

SUPPORTS_INIT_PRIORITY

If nonzero, the C++ init_priority attribute is supported and the compiler should emit instructions to control the order of initialization of objects. If zero, the compiler will issue an error message upon encountering an init_priority attribute.

bool TARGET_HAVE_CTORS_DTORS

This value is true if the target supports some “native” method of collecting constructors and destructors to be run at startup and exit. It is false if we must use collect2.

void TARGET_ASM_CONSTRUCTOR (rtx symbol, int priority)

If defined, a function that outputs assembler code to arrange to call the function referenced by symbol at initialization time.

Assume that symbol is a SYMBOL_REF for a function taking no arguments and with no return value. If the target supports initialization priorities, priority is a value between 0 and MAX_INIT_PRIORITY; otherwise it must be DEFAULT_INIT_PRIORITY.

If this macro is not defined by the target, a suitable default will be chosen if (1) the target supports arbitrary section names, (2) the target defines CTORS_SECTION_ASM_OP, or (3) USE_COLLECT2 is not defined.

void TARGET_ASM_DESTRUCTOR (rtx symbol, int priority)

This is like TARGET_ASM_CONSTRUCTOR but used for termination functions rather than initialization functions.

If TARGET_HAVE_CTORS_DTORS is true, the initialization routine generated for the generated object file will have static linkage.

If your system uses collect2 as the means of processing constructors, then that program normally uses nm to scan an object file for constructor functions to be called.

On certain kinds of systems, you can define this macro to make collect2 work faster (and, in some cases, make it work at all):

OBJECT_FORMAT_COFF

Define this macro if the system uses COFF (Common Object File Format) object files, so that collect2 can assume this format and scan object files directly for dynamic constructor/destructor functions.

This macro is effective only in a native compiler; collect2 as part of a cross compiler always uses nm for the target machine.

REAL_NM_FILE_NAME

Define this macro as a C string constant containing the file name to use to execute nm. The default is to search the path normally for nm.
NM_FLAGS

`collect2` calls `nm` to scan object files for static constructors and destructors and LTO info. By default, `-n` is passed. Define `NM_FLAGS` to a C string constant if other options are needed to get the same output format as GNU `nm -n` produces.

If your system supports shared libraries and has a program to list the dynamic dependencies of a given library or executable, you can define these macros to enable support for running initialization and termination functions in shared libraries:

LDD_SUFFIX

Define this macro to a C string constant containing the name of the program which lists dynamic dependencies, like `ldd` under SunOS 4.

PARSE_LDD_OUTPUT (ptr)

Define this macro to be C code that extracts filenames from the output of the program denoted by `LDD_SUFFIX`. `ptr` is a variable of type `char *` that points to the beginning of a line of output from `LDD_SUFFIX`. If the line lists a dynamic dependency, the code must advance `ptr` to the beginning of the filename on that line. Otherwise, it must set `ptr` to `NULL`.

SHLIB_SUFFIX

Define this macro to a C string constant containing the default shared library extension of the target (e.g., `".so"`). `collect2` strips version information after this suffix when generating global constructor and destructor names. This define is only needed on targets that use `collect2` to process constructors and destructors.

17.21.7 Output of Assembler Instructions

This describes assembler instruction output.

REGISTER_NAMES

A C initializer containing the assembler’s names for the machine registers, each one as a C string constant. This is what translates register numbers in the compiler into assembler language.

ADDITIONAL_REGISTER_NAMES

If defined, a C initializer for an array of structures containing a name and a register number. This macro defines additional names for hard registers, thus allowing the `asm` option in declarations to refer to registers using alternate names.

OVERLAPPING_REGISTER_NAMES

If defined, a C initializer for an array of structures containing a name, a register number and a count of the number of consecutive machine registers the name overlaps. This macro defines additional names for hard registers, thus allowing the `asm` option in declarations to refer to registers using alternate names. Unlike `ADDITIONAL_REGISTER_NAMES`, this macro should be used when the register name implies multiple underlying registers.

This macro should be used when it is important that a clobber in an `asm` statement clobbers all the underlying values implied by the register name. For example, on ARM, clobbering the double-precision VFP register “d0” implies clobbering both single-precision registers “s0” and “s1”.


ASM_OUTPUT_OPCODE (stream, ptr)  

Define this macro if you are using an unusual assembler that requires different names for the machine instructions.

The definition is a C statement or statements which output an assembler instruction opcode to the stdio stream stream. The macro-operand ptr is a variable of type char * which points to the opcode name in its “internal” form—the form that is written in the machine description. The definition should output the opcode name to stream, performing any translation you desire, and increment the variable ptr to point at the end of the opcode so that it will not be output twice.

In fact, your macro definition may process less than the entire opcode name, or more than the opcode name; but if you want to process text that includes ‘%’-sequences to substitute operands, you must take care of the substitution yourself. Just be sure to increment ptr over whatever text should not be output normally.

If you need to look at the operand values, they can be found as the elements of recog_data.operand.

If the macro definition does nothing, the instruction is output in the usual way.

FINAL_PRESCAN_INSN (insn, opvec, noperands)  

If defined, a C statement to be executed just prior to the output of assembler code for insn, to modify the extracted operands so they will be output differently.

Here the argument opvec is the vector containing the operands extracted from insn, and noperands is the number of elements of the vector which contain meaningful data for this insn. The contents of this vector are what will be used to convert the insn template into assembler code, so you can change the assembler output by changing the contents of the vector.

This macro is useful when various assembler syntaxes share a single file of instruction patterns; by defining this macro differently, you can cause a large class of instructions to be output differently (such as with rearranged operands). Naturally, variations in assembler syntax affecting individual insn patterns ought to be handled by writing conditional output routines in those patterns.

If this macro is not defined, it is equivalent to a null statement.

void TARGET_ASM_FINAL_POSTSCAN_INSN (FILE *file, rtx insn, rtx *opvec, int noperands)  

If defined, this target hook is a function which is executed just after the output of assembler code for insn, to modify the mode of the assembler if necessary.

Here the argument opvec is the vector containing the operands extracted from insn, and noperands is the number of elements of the vector which contain meaningful data for this insn. The contents of this vector are what was used to convert the insn template into assembler code, so you can change the assembler mode by checking the contents of the vector.

PRINT_OPERAND (stream, x, code)  

A C compound statement to output to stdio stream stream the assembler syntax for an instruction operand x. x is an RTL expression.
code is a value that can be used to specify one of several ways of printing the operand. It is used when identical operands must be printed differently depending on the context. code comes from the ‘%’ specification that was used to request printing of the operand. If the specification was just ‘%digit’ then code is 0; if the specification was ‘%ltr digit’ then code is the ASCII code for ltr.

If x is a register, this macro should print the register’s name. The names can be found in an array reg_names whose type is char *[]. reg_names is initialized from REGISTER_NAMES.

When the machine description has a specification ‘%punct’ (a ‘%’ followed by a punctuation character), this macro is called with a null pointer for x and the punctuation character for code.

PRINT_OPERAND_PUNCT_VALID_P (code) [Macro]
A C expression which evaluates to true if code is a valid punctuation character for use in the PRINT_OPERAND macro. If PRINT_OPERAND_PUNCT_VALID_P is not defined, it means that no punctuation characters (except for the standard one, ‘%’) are used in this way.

PRINT_OPERAND_ADDRESS (stream, x) [Macro]
A C compound statement to output to stdio stream stream the assembler syntax for an instruction operand that is a memory reference whose address is x. x is an RTL expression.

On some machines, the syntax for a symbolic address depends on the section that the address refers to. On these machines, define the hook TARGET_ENCODE_SECTION_INFO to store the information into the symbol_ref, and then check for it here. See Section 17.21 [Assembler Format], page 497.

DBR_OUTPUT_SEQEND (file) [Macro]
A C statement, to be executed after all slot-filler instructions have been output. If necessary, call dbr_sequence_length to determine the number of slots filled in a sequence (zero if not currently outputting a sequence), to decide how many no-ops to output, or whatever.

Don’t define this macro if it has nothing to do, but it is helpful in reading assembly output if the extent of the delay sequence is made explicit (e.g. with white space).

Note that output routines for instructions with delay slots must be prepared to deal with not being output as part of a sequence (i.e. when the scheduling pass is not run, or when no slot fillers could be found.) The variable final_sequence is null when not processing a sequence, otherwise it contains the sequence rtx being output.

REGISTER_PREFIX [Macro]
LOCAL_LABEL_PREFIX [Macro]
USER_LABEL_PREFIX [Macro]
IMMEDIATE_PREFIX [Macro]
If defined, C string expressions to be used for the ‘%R’, ‘%L’, ‘%U’, and ‘%I’ options of asm_fprintf (see ‘final.c’). These are useful when a single ‘md’ file must support multiple assembler formats. In that case, the various ‘tm.h’ files can define these macros differently.
ASM_PRINTF_EXTENSIONS (file, argptr, format)  [Macro]
If defined this macro should expand to a series of case statements which will be parsed inside the switch statement of the asm_fprintf function. This allows targets to define extra printf formats which may be useful when generating their assembler statements. Note that uppercase letters are reserved for future generic extensions to asm_fprintf, and so are not available to target specific code. The output file is given by the parameter file. The varargs input pointer is argptr and the rest of the format string, starting the character after the one that is being switched upon, is pointed to by format.

ASSEMBLER_DIALECT  [Macro]
If your target supports multiple dialects of assembler language (such as different opcodes), define this macro as a C expression that gives the numeric index of the assembler language dialect to use, with zero as the first variant.
If this macro is defined, you may use constructs of the form
‘{option0|option1|option2...}’
in the output templates of patterns (see Section 16.5 [Output Template], page 295) or in the first argument of asm_fprintf. This construct outputs ‘option0’, ‘option1’, ‘option2’, etc., if the value of ASSEMBLER_DIALECT is zero, one, two, etc. Any special characters within these strings retain their usual meaning. If there are fewer alternatives within the braces than the value of ASSEMBLER_DIALECT, the construct outputs nothing.
If you do not define this macro, the characters ‘{’, ‘|’ and ‘}’ do not have any special meaning when used in templates or operands to asm_fprintf.
Define the macros REGISTER_PREFIX, LOCAL_LABEL_PREFIX, USER_LABEL_PREFIX and IMMEDIATE_PREFIX if you can express the variations in assembler language syntax with that mechanism. Define ASSEMBLER_DIALECT and use the ‘{option0|option1}’ syntax if the syntax variant are larger and involve such things as different opcodes or operand order.

ASM_OUTPUT_REG_PUSH (stream, regno)  [Macro]
A C expression to output to stream some assembler code which will push hard register number regno onto the stack. The code need not be optimal, since this macro is used only when profiling.

ASM_OUTPUT_REG_POP (stream, regno)  [Macro]
A C expression to output to stream some assembler code which will pop hard register number regno off of the stack. The code need not be optimal, since this macro is used only when profiling.

17.21.8 Output of Dispatch Tables
This concerns dispatch tables.

ASM_OUTPUT_ADDR_DIFF_ELT (stream, body, value, rel)  [Macro]
A C statement to output to the stdio stream stream an assembler pseudo-instruction to generate a difference between two labels. value and rel are the numbers of two internal labels. The definitions of these labels are output using (*targetm.asm_out.internal_label), and they must be printed in the same way here. For example,
fprintf (stream, "\t.word L%d-L%d\n", value, rel)

You must provide this macro on machines where the addresses in a dispatch table are relative to the table’s own address. If defined, GCC will also use this macro on all machines when producing PIC. body is the body of the ADDR_DIFF_VEC; it is provided so that the mode and flags can be read.

**ASM_OUTPUT_ADDR_VEC_ELT (stream, value)**

This macro should be provided on machines where the addresses in a dispatch table are absolute.

The definition should be a C statement to output to the stdio stream stream an assembler pseudo-instruction to generate a reference to a label. value is the number of an internal label whose definition is output using (*targetm.asm_out.internal_label). For example,

```c
fprintf (stream, "\t.word L%d\n", value)
```

**ASM_OUTPUT_CASE_LABEL (stream, prefix, num, table)**

Define this if the label before a jump-table needs to be output specially. The first three arguments are the same as for (*targetm.asm_out.internal_label); the fourth argument is the jump-table which follows (a jump_insn containing an addr_vec or addr_diff_vec).

This feature is used on system V to output a swbeg statement for the table.

If this macro is not defined, these labels are output with (*targetm.asm_out.internal_label).

**ASM_OUTPUT_CASE_END (stream, num, table)**

Define this if something special must be output at the end of a jump-table. The definition should be a C statement to be executed after the assembler code for the table is written. It should write the appropriate code to stdio stream stream. The argument table is the jump-table insn, and num is the label-number of the preceding label.

If this macro is not defined, nothing special is output at the end of the jump-table.

**TARGET_ASM_EMIT_UNWIND_LABEL (FILE *stream, tree decl, int for_eh, int empty)**

This target hook emits a label at the beginning of each FDE. It should be defined on targets where FDEs need special labels, and it should write the appropriate label, for the FDE associated with the function declaration decl, to the stdio stream stream. The third argument, for_eh, is a boolean: true if this is for an exception table. The fourth argument, empty, is a boolean: true if this is a placeholder label for an omitted FDE.

The default is that FDEs are not given nonlocal labels.

**TARGET_ASM_EMIT_EXCEPT_TABLE_LABEL (FILE *stream)**

This target hook emits a label at the beginning of the exception table. It should be defined on targets where it is desirable for the table to be broken up according to function.

The default is that no label is emitted.
Chapter 17: Target Description Macros and Functions

void TARGET_ASM_EMIT_EXCEPT_PERSONALITY (rtx personality)  [Target Hook]
If the target implements TARGET_ASM_UNWIND_EMIT, this hook may be used to emit
a directive to install a personality hook into the unwind info. This hook should not
be used if dwarf2 unwind info is used.

void TARGET_ASM_UNWIND_EMIT (FILE *stream, rtx insn)  [Target Hook]
This target hook emits assembly directives required to unwind the given instruction.
This is only used when TARGET_EXCEPTION_UNWIND_INFO returns UI_TARGET.

bool TARGET_ASM_UNWIND_EMIT_BEFORE_INSN  [Target Hook]
True if the TARGET_ASM_UNWIND_EMIT hook should be called before the assembly for
insn has been emitted, false if the hook should be called afterward.

17.21.9 Assembler Commands for Exception Regions
This describes commands marking the start and the end of an exception region.

EH_FRAME_SECTION_NAME  [Macro]
If defined, a C string constant for the name of the section containing exception han-
dling frame unwind information. If not defined, GCC will provide a default definition
if the target supports named sections. ‘crtstuff.c’ uses this macro to switch to the
appropriate section.

You should define this symbol if your target supports DWARF 2 frame unwind infor-
mation and the default definition does not work.

EH_FRAME_IN_DATA_SECTION  [Macro]
If defined, DWARF 2 frame unwind information will be placed in the data section even
though the target supports named sections. This might be necessary, for instance, if
the system linker does garbage collection and sections cannot be marked as not to be
collected.

Do not define this macro unless TARGET_ASM_NAMED_SECTION is also defined.

EH_TABLES_CAN_BE_READ_ONLY  [Macro]
Define this macro to 1 if your target is such that no frame unwind information en-
coding used with non-PIC code will ever require a runtime relocation, but the linker
may not support merging read-only and read-write sections into a single read-write
section.

MASK_RETURN_ADDR  [Macro]
An rtx used to mask the return address found via RETURN_ADDR_RTX, so that it does
not contain any extraneous set bits in it.

DWARF2_UNWIND_INFO  [Macro]
Define this macro to 0 if your target supports DWARF 2 frame unwind information,
but it does not yet work with exception handling. Otherwise, if your target supports
this information (if it defines INCOMING_RETURN_ADDR_RTX and either UNALIGNED-
INT_ASM_OP or OBJECT_FORMAT_ELF), GCC will provide a default definition of 1.
enum unwind_info_type TARGET_EXCEPT_UNWIND_INFO (struct gcc_options *opts)

This hook defines the mechanism that will be used for exception handling by the target. If the target has ABI specified unwind tables, the hook should return UI_TARGET. If the target is to use the setjmp/longjmp-based exception handling scheme, the hook should return UI_SJLJ. If the target supports DWARF 2 frame unwind information, the hook should return UI_DWARF2.

A target may, if exceptions are disabled, choose to return UI_NONE. This may end up simplifying other parts of target-specific code. The default implementation of this hook never returns UI_NONE.

Note that the value returned by this hook should be constant. It should not depend on anything except the command-line switches described by opts. In particular, the setting UI_SJLJ must be fixed at compiler start-up as C pre-processor macros and builtin functions related to exception handling are set up depending on this setting.

The default implementation of the hook first honors the '--enable-sjlj-exceptions' configure option, then DWARF2_UNWIND_INFO, and finally defaults to UI_SJLJ. If DWARF2_UNWIND_INFO depends on command-line options, the target must define this hook so that opts is used correctly.

bool TARGET_UNWIND_TABLES_DEFAULT

This variable should be set to true if the target ABI requires unwinding tables even when exceptions are not used. It must not be modified by command-line option processing.

DONT_USE_BUILTIN_SETJMP

Define this macro to 1 if the setjmp/longjmp-based scheme should use the setjmp/longjmp functions from the C library instead of the __builtin_setjmp/__builtin_longjmp machinery.

DWARF_CIE_DATA_ALIGNMENT

This macro need only be defined if the target might save registers in the function prologue at an offset to the stack pointer that is not aligned to UNITS_PER_WORD. The definition should be the negative minimum alignment if STACK_GROWS_DOWNWARD is defined, and the positive minimum alignment otherwise. See Section 17.22.5 [SDB and DWARF], page 529. Only applicable if the target supports DWARF 2 frame unwind information.

bool TARGET_TERMINATE_DW2_EH_FRAME_INFO

Contains the value true if the target should add a zero word onto the end of a Dwarf-2 frame info section when used for exception handling. Default value is false if EH_FRAME_SECTION_NAME is defined, and true otherwise.

rtx TARGET_DWARF_REGISTER_SPAN (rtx reg)

Given a register, this hook should return a parallel of registers to represent where to find the register pieces. Define this hook if the register and its mode are represented in Dwarf in non-contiguous locations, or if the register should be represented in more than one register in Dwarf. Otherwise, this hook should return NULL_RTX. If not defined, the default is to return NULL_RTX.
void TARGET_INIT_DWARF_REG_SIZES_EXTRA (tree address)  [Target Hook]
If some registers are represented in Dwarf-2 unwind information in multiple pieces, define this hook to fill in information about the sizes of those pieces in the table used by the unwinder at runtime. It will be called by expand_builtin_init_dwarf_reg_sizes after filling in a single size corresponding to each hard register; address is the address of the table.

bool TARGET_ASM_TTYPE (rtx sym)  [Target Hook]
This hook is used to output a reference from a frame unwinding table to the type_info object identified by sym. It should return true if the reference was output. Returning false will cause the reference to be output using the normal Dwarf2 routines.

bool TARGET_ARM_EABI_UNWINDER  [Target Hook]
This flag should be set to true on targets that use an ARM EABI based unwinding library, and false on other targets. This effects the format of unwinding tables, and how the unwinder in entered after running a cleanup. The default is false.

17.21.10 Assembler Commands for Alignment
This describes commands for alignment.

JUMP_ALIGN (label)  [Macro]
The alignment (log base 2) to put in front of label, which is a common destination of jumps and has no fallthru incoming edge.
This macro need not be defined if you don’t want any special alignment to be done at such a time. Most machine descriptions do not currently define the macro. Unless it’s necessary to inspect the label parameter, it is better to set the variable align_jumps in the target’s TARGET_OPTION_OVERRIDE. Otherwise, you should try to honor the user’s selection in align_jumps in a JUMP_ALIGN implementation.

int TARGET_ASM_JUMP_ALIGN_MAX_SKIP (rtx label)  [Target Hook]
The maximum number of bytes to skip before label when applying JUMP_ALIGN. This works only if ASM_OUTPUT_MAX_SKIP_ALIGN is defined.

LABEL_ALIGN_AFTER_BARRIER (label)  [Macro]
The alignment (log base 2) to put in front of label, which follows a BARRIER.
This macro need not be defined if you don’t want any special alignment to be done at such a time. Most machine descriptions do not currently define the macro.

int TARGET_ASM_LABEL_ALIGN_AFTER_BARRIER_MAX_SKIP (rtx label)  [Target Hook]
The maximum number of bytes to skip before label when applying LABEL_ALIGN_AFTER_BARRIER. This works only if ASM_OUTPUT_MAX_SKIP_ALIGN is defined.

LOOP_ALIGN (label)  [Macro]
The alignment (log base 2) to put in front of label, which follows a NOTE_INSN LOOP_BEG note.
This macro need not be defined if you don’t want any special alignment to be done at such a time. Most machine descriptions do not currently define the macro.
Unless it’s necessary to inspect the *label* parameter, it is better to set the variable *align_loops* in the target’s `TARGET_OPTION_OVERRIDE`. Otherwise, you should try to honor the user’s selection in *align_loops* in a `LOOP_ALIGN` implementation.

```c
int TARGET_ASM_LOOP_ALIGN_MAX_SKIP (rtx label) [Target Hook]
The maximum number of bytes to skip when applying `LOOP_ALIGN` to *label*. This works only if `ASM_OUTPUT_MAX_SKIP_ALIGN` is defined.
```

```c
LABEL_ALIGN (label) [Macro]
The alignment (log base 2) to put in front of *label*. If `LABEL_ALIGN_AFTER_BARRIER` / `LOOP_ALIGN` specify a different alignment, the maximum of the specified values is used.

Unless it’s necessary to inspect the *label* parameter, it is better to set the variable *align_labels* in the target’s `TARGET_OPTION_OVERRIDE`. Otherwise, you should try to honor the user’s selection in *align_labels* in a `LABEL_ALIGN` implementation.

```c
int TARGET_ASM_LABEL_ALIGN_MAX_SKIP (rtx label) [Target Hook]
The maximum number of bytes to skip when applying `LABEL_ALIGN` to *label*. This works only if `ASM_OUTPUT_MAX_SKIP_ALIGN` is defined.
```

```c
ASM_OUTPUT_SKIP (stream, nbytes) [Macro]
A C statement to output to the stdio stream *stream* an assembler instruction to advance the location counter by *nbytes* bytes. Those bytes should be zero when loaded. *nbytes* will be a C expression of type `unsigned HOST_WIDE_INT`.
```

```c
ASM_NO_SKIP_IN_TEXT [Macro]
Define this macro if `ASM_OUTPUT_SKIP` should not be used in the text section because it fails to put zeros in the bytes that are skipped. This is true on many Unix systems, where the pseudo–op to skip bytes produces no-op instructions rather than zeros when used in the text section.
```

```c
ASM_OUTPUT_ALIGN (stream, power) [Macro]
A C statement to output to the stdio stream *stream* an assembler command to advance the location counter to a multiple of 2 to the *power* bytes. *power* will be a C expression of type `int`.
```

```c
ASM_OUTPUT_ALIGN_WITH_NOP (stream, power) [Macro]
Like `ASM_OUTPUT_ALIGN`, except that the “nop” instruction is used for padding, if necessary.
```

```c
ASM_OUTPUT_MAX_SKIP_ALIGN (stream, power, max_skip) [Macro]
A C statement to output to the stdio stream *stream* an assembler command to advance the location counter to a multiple of 2 to the *power* bytes, but only if *max_skip* or fewer bytes are needed to satisfy the alignment request. *power* and *max_skip* will be a C expression of type `int`.
```
17.22 Controlling Debugging Information Format

This describes how to specify debugging information.

17.22.1 Macros Affecting All Debugging Formats

These macros affect all debugging formats.

**DBX_REGISTER_NUMBER (regno)**

A C expression that returns the DBX register number for the compiler register number `regno`. In the default macro provided, the value of this expression will be `regno` itself. But sometimes there are some registers that the compiler knows about and DBX does not, or vice versa. In such cases, some register may need to have one number in the compiler and another for DBX.

If two registers have consecutive numbers inside GCC, and they can be used as a pair to hold a multiword value, then they *must* have consecutive numbers after renumbering with `DBX_REGISTER_NUMBER`. Otherwise, debuggers will be unable to access such a pair, because they expect register pairs to be consecutive in their own numbering scheme.

If you find yourself defining `DBX_REGISTER_NUMBER` in a way that does not preserve register pairs, then what you must do instead is redefine the actual register numbering scheme.

**DEBUGGER_AUTO_OFFSET (x)**

A C expression that returns the integer offset value for an automatic variable having address `x` (an RTL expression). The default computation assumes that `x` is based on the frame-pointer and gives the offset from the frame-pointer. This is required for targets that produce debugging output for DBX or COFF-style debugging output for SDB and allow the frame-pointer to be eliminated when the `-g` options is used.

**DEBUGGER_ARG_OFFSET (offset, x)**

A C expression that returns the integer offset value for an argument having address `x` (an RTL expression). The nominal offset is `offset`.

**PREFERRED_DEBUGGING_TYPE**

A C expression that returns the type of debugging output GCC should produce when the user specifies just `-g`. Define this if you have arranged for GCC to support more than one format of debugging output. Currently, the allowable values are `DBX_DEBUG`, `SDB_DEBUG`, `DWARF_DEBUG`, `DWARF2_DEBUG`, `XCOFF_DEBUG`, `VMS_DEBUG`, and `VMS_AND_DWARF2_DEBUG`.

When the user specifies `-ggdb`, GCC normally also uses the value of this macro to select the debugging output format, but with two exceptions. If `DWARF2_DEBUGGING_INFO` is defined, GCC uses the value `DWARF2_DEBUG`. Otherwise, if `DBX_DEBUGGING_INFO` is defined, GCC uses `DBX_DEBUG`.

The value of this macro only affects the default debugging output; the user can always get a specific type of output by using `'-gstabs'`, `'-gcoff'`, `'-gdwarf-2'`, `'-gxcoff'`, or `'-gvms'`. 
17.22.2 Specific Options for DBX Output

These are specific options for DBX output.

**DBX_DEBUGGING_INFO** [Macro]
Define this macro if GCC should produce debugging output for DBX in response to the ‘-g’ option.

**XCOFF_DEBUGGING_INFO** [Macro]
Define this macro if GCC should produce XCOFF format debugging output in response to the ‘-g’ option. This is a variant of DBX format.

**DEFAULT_GDB_EXTENSIONS** [Macro]
Define this macro to control whether GCC should by default generate GDB’s extended version of DBX debugging information (assuming DBX-format debugging information is enabled at all). If you don’t define the macro, the default is 1: always generate the extended information if there is any occasion to.

**DEBUG_SYMS_TEXT** [Macro]
Define this macro if all .stabs commands should be output while in the text section.

**ASM_STABS_OP** [Macro]
A C string constant, including spacing, naming the assembler pseudo op to use instead of "\t.stabs\t" to define an ordinary debugging symbol. If you don’t define this macro, "\t.stabs\t" is used. This macro applies only to DBX debugging information format.

**ASM_STABD_OP** [Macro]
A C string constant, including spacing, naming the assembler pseudo op to use instead of "\t.stabd\t" to define a debugging symbol whose value is the current location. If you don’t define this macro, "\t.stabd\t" is used. This macro applies only to DBX debugging information format.

**ASM_STABN_OP** [Macro]
A C string constant, including spacing, naming the assembler pseudo op to use instead of "\t.stabn\t" to define a debugging symbol with no name. If you don’t define this macro, "\t.stabn\t" is used. This macro applies only to DBX debugging information format.

**DBX_NO_XREFS** [Macro]
Define this macro if DBX on your system does not support the construct ‘\textit{xstagname}’. On some systems, this construct is used to describe a forward reference to a structure named \textit{tagname}. On other systems, this construct is not supported at all.

**DBX_CONTIN_LENGTH** [Macro]
A symbol name in DBX-format debugging information is normally continued (split into two separate .stabs directives) when it exceeds a certain length (by default, 80 characters). On some operating systems, DBX requires this splitting; on others, splitting must not be done. You can inhibit splitting by defining this macro with the value zero. You can override the default splitting-length by defining this macro as an expression for the length you desire.
Chapter 17: Target Description Macros and Functions

DBX_CONTIN_CHAR

Normally continuation is indicated by adding a `\' character to the end of a .stabs string when a continuation follows. To use a different character instead, define this macro as a character constant for the character you want to use. Do not define this macro if backslash is correct for your system.

DBX_STATIC_STAB_DATA_SECTION

Define this macro if it is necessary to go to the data section before outputting the `.stabs' pseudo-op for a non-global static variable.

DBX_TYPE_DECL_STABS_CODE

The value to use in the “code” field of the .stabs directive for a typedef. The default is N_LSYM.

DBX_STATIC_CONST_VAR_CODE

The value to use in the “code” field of the .stabs directive for a static variable located in the text section. DBX format does not provide any “right” way to do this. The default is N_FUN.

DBX_REGPARM_STABS_CODE

The value to use in the “code” field of the .stabs directive for a parameter passed in registers. DBX format does not provide any “right” way to do this. The default is N_RSYM.

DBX_REGPARM_STABS_LETTER

The letter to use in DBX symbol data to identify a symbol as a parameter passed in registers. DBX format does not customarily provide any way to do this. The default is ’P’.

DBX_FUNCTION_FIRST

Define this macro if the DBX information for a function and its arguments should precede the assembler code for the function. Normally, in DBX format, the debugging information entirely follows the assembler code.

DBX_BLOCKS_FUNCTION_RELATIVE

Define this macro, with value 1, if the value of a symbol describing the scope of a block (N_LBRAC or N_RBRAC) should be relative to the start of the enclosing function. Normally, GCC uses an absolute address.

DBX_LINES_FUNCTION_RELATIVE

Define this macro, with value 1, if the value of a symbol indicating the current line number (N_SLINE) should be relative to the start of the enclosing function. Normally, GCC uses an absolute address.

DBX_USE_BINCL

Define this macro if GCC should generate N_BINCL and N_EINCL stabs for included header files, as on Sun systems. This macro also directs GCC to output a type number as a pair of a file number and a type number within the file. Normally, GCC does not generate N_BINCL or N_EINCL stabs, and it outputs a single number for a type number.
17.22.3 Open-Ended Hooks for DBX Format

These are hooks for DBX format.

DBX_OUTPUT_LBRAC (stream, name)  [Macro]
Define this macro to say how to output to stream the debugging information for the start of a scope level for variable names. The argument name is the name of an assembler symbol (for use with assemble_name) whose value is the address where the scope begins.

DBX_OUTPUT_RBRAC (stream, name)  [Macro]
Like DBX_OUTPUT_LBRAC, but for the end of a scope level.

DBX_OUTPUT_NFUN (stream, lscope_label, decl)  [Macro]
Define this macro if the target machine requires special handling to output an N_FUN entry for the function decl.

DBX_OUTPUT_SOURCE_LINE (stream, line, counter)  [Macro]
A C statement to output DBX debugging information before code for line number line of the current source file to the stdio stream stream. counter is the number of time the macro was invoked, including the current invocation; it is intended to generate unique labels in the assembly output.

This macro should not be defined if the default output is correct, or if it can be made correct by defining DBX_LINES_FUNCTION_RELATIVE.

NO_DBX_FUNCTION_END  [Macro]
Some stabs encapsulation formats (in particular ECOFF), cannot handle the .stabs ",,N_FUN,,0,0,Lscope-function-1 gdb dbx extension construct. On those machines, define this macro to turn this feature off without disturbing the rest of the gdb extensions.

NO_DBX_BNSYM_ENSYM  [Macro]
Some assemblers cannot handle the .stabd BNSYM/ENSYM,0,0 gdb dbx extension construct. On those machines, define this macro to turn this feature off without disturbing the rest of the gdb extensions.

17.22.4 File Names in DBX Format

This describes file names in DBX format.

DBX_OUTPUT_MAIN_SOURCE_FILENAME (stream, name)  [Macro]
A C statement to output DBX debugging information to the stdio stream stream, which indicates that file name is the main source file—the file specified as the input file for compilation. This macro is called only once, at the beginning of compilation.

This macro need not be defined if the standard form of output for DBX debugging information is appropriate.

It may be necessary to refer to a label equal to the beginning of the text section. You can use 'assemble_name (stream, ltext_label_name)' to do so. If you do this, you must also set the variable used_ltext_label_name to true.
NO_DBX_MAIN_SOURCE_DIRECTORY  [Macro]
Define this macro, with value 1, if GCC should not emit an indication of the current directory for compilation and current source language at the beginning of the file.

NO_DBX_GCC_MARKER  [Macro]
Define this macro, with value 1, if GCC should not emit an indication that this object file was compiled by GCC. The default is to emit an N_OPT stab at the beginning of every source file, with ‘gcc2_compiled.’ for the string and value 0.

DBX_OUTPUT_MAIN_SOURCE_FILE_END  [Macro]
A C statement to output DBX debugging information at the end of compilation of the main source file name. Output should be written to the stdio stream stream.
If you don’t define this macro, nothing special is output at the end of compilation, which is correct for most machines.

DBX_OUTPUT_NULL_N_SO_AT_MAIN_SOURCE_FILE_END  [Macro]
Define this macro instead of defining DBX_OUTPUT_MAIN_SOURCE_FILE_END, if what needs to be output at the end of compilation is an N_SO stab with an empty string, whose value is the highest absolute text address in the file.

17.22.5 Macros for SDB and DWARF Output
Here are macros for SDB and DWARF output.

SDB_DEBUGGING_INFO  [Macro]
Define this macro if GCC should produce COFF-style debugging output for SDB in response to the ‘-g’ option.

DWARF2_DEBUGGING_INFO  [Macro]
Define this macro if GCC should produce dwarf version 2 format debugging output in response to the ‘-g’ option.

int TARGET_DWARF_CALLING_CONVENTION (const_tree function)  [Target Hook]
Define this to enable the dwarf attribute DW_AT_calling_convention to be emitted for each function. Instead of an integer return the enum value for the DW_CC_ tag.

To support optional call frame debugging information, you must also define INCOMING_RETURN_ADDR RTX and either set RTX_FRAME_RELATED_P on the prologue insns if you use RTL for the prologue, or call dwarf2out_def_cfa and dwarf2out_reg_save as appropriate from TARGET_ASM_FUNCTION_PROLOGUE if you don’t.

DWARF2_FRAME_INFO  [Macro]
Define this macro to a nonzero value if GCC should always output Dwarf 2 frame information. If TARGET_EXCEPT_UNWIND_INFO (see Section 17.21.9 [Exception Region Output], page 521) returns UI_DWARF2, and exceptions are enabled, GCC will output this information not matter how you define DWARF2_FRAME_INFO.
enum unwind_info_type TARGET_DEBUG_UNWIND_INFO (void) [Target Hook]

This hook defines the mechanism that will be used for describing frame unwind information to the debugger. Normally the hook will return UI_DWARF2 if DWARF 2 debug information is enabled, and return UI_NONE otherwise.

A target may return UI_DWARF2 even when DWARF 2 debug information is disabled in order to always output DWARF 2 frame information.

A target may return UI_TARGET if it has ABI specified unwind tables. This will suppress generation of the normal debug frame unwind information.

DWARF2_ASM_LINE_DEBUG_INFO [Macro]

Define this macro to be a nonzero value if the assembler can generate Dwarf 2 line debug info sections. This will result in much more compact line number tables, and hence is desirable if it works.

bool TARGET_WANT_DEBUG_PUB_SECTIONS [Target Hook]

True if the .debug_pubtypes and .debug_pubnames sections should be emitted. These sections are not used on most platforms, and in particular GDB does not use them.

bool TARGET_DELAY_SCHED2 [Target Hook]

True if sched2 is not to be run at its normal place. This usually means it will be run as part of machine-specific reorg.

bool TARGET_DELAY_VARTRACK [Target Hook]

True if vartrack is not to be run at its normal place. This usually means it will be run as part of machine-specific reorg.

ASM_OUTPUT_DWARF_DELTA (stream, size, label1, label2) [Macro]

A C statement to issue assembly directives that create a difference lab1 minus lab2, using an integer of the given size.

ASM_OUTPUT_DWARF_VMS_DELTA (stream, size, label1, label2) [Macro]

A C statement to issue assembly directives that create a difference between the two given labels in system defined units, e.g. instruction slots on IA64 VMS, using an integer of the given size.

ASM_OUTPUT_DWARF_OFFSET (stream, size, label, section) [Macro]

A C statement to issue assembly directives that create a section-relative reference to the given label, using an integer of the given size. The label is known to be defined in the given section.

ASM_OUTPUT_DWARF_PCREL (stream, size, label) [Macro]

A C statement to issue assembly directives that create a self-relative reference to the given label, using an integer of the given size.

ASM_OUTPUT_DWARF_TABLE_REF (label) [Macro]

A C statement to issue assembly directives that create a reference to the DWARF table identifier label from the current section. This is used on some systems to avoid garbage collecting a DWARF table which is referenced by a function.
Chapter 17: Target Description Macros and Functions

void TARGET_ASM_OUTPUT_DWARF_DTPREL (FILE *file, int size, [Target Hook]
       rtx x)
   If defined, this target hook is a function which outputs a DTP-relative reference to
   the given TLS symbol of the specified size.

PUT_SDB...
   [Macro]
   Define these macros to override the assembler syntax for the special SDB assembler
directives. See ‘sdbout.c’ for a list of these macros and their arguments. If the
standard syntax is used, you need not define them yourself.

SDB_DELIM
   [Macro]
   Some assemblers do not support a semicolon as a delimiter, even between SDB as-
sembler directives. In that case, define this macro to be the delimiter to use (usually
‘\n’). It is not necessary to define a new set of PUT_SDB_<op> macros if this is the only
change required.

SDB_ALLOW_UNKNOWN_REFERENCES
   [Macro]
   Define this macro to allow references to unknown structure, union, or enumeration
tags to be emitted. Standard COFF does not allow handling of unknown references,
MIPS ECOFF has support for it.

SDB_ALLOW_FORWARD_REFERENCES
   [Macro]
   Define this macro to allow references to structure, union, or enumeration tags that
have not yet been seen to be handled. Some assemblers choke if forward tags are
used, while some require it.

SDB_OUTPUT_SOURCE_LINE (stream, line)
   [Macro]
   A C statement to output SDB debugging information before code for line number
line of the current source file to thestdio stream stream. The default is to emit an
.ln directive.

17.22.6 Macros for VMS Debug Format

Here are macros for VMS debug format.

VMS_DEBUGGING_INFO
   [Macro]
   Define this macro if GCC should produce debugging output for VMS in response to
the ‘-g’ option. The default behavior for VMS is to generate minimal debug info for a
traceback in the absence of ‘-g’ unless explicitly overridden with ‘-g0’. This behavior
is controlled by TARGET_OPTION_OPTIMIZATION and TARGET_OPTION_OVERRIDE.

17.23 Cross Compilation and Floating Point

While all modern machines use twos-complement representation for integers, there are a
variety of representations for floating point numbers. This means that in a cross-compiler
the representation of floating point numbers in the compiled program may be different from
that used in the machine doing the compilation.

Because different representation systems may offer different amounts of range and pre-
cision, all floating point constants must be represented in the target machine’s format.
Therefore, the cross compiler cannot safely use the host machine’s floating point arithmetic; it must emulate the target’s arithmetic. To ensure consistency, GCC always uses emulation to work with floating point values, even when the host and target floating point formats are identical.

The following macros are provided by ‘real.h’ for the compiler to use. All parts of the compiler which generate or optimize floating-point calculations must use these macros. They may evaluate their operands more than once, so operands must not have side effects.

**REAL_VALUE_TYPE**

The C data type to be used to hold a floating point value in the target machine’s format. Typically this is a `struct` containing an array of `HOST_WIDE_INT`, but all code should treat it as an opaque quantity.

**int REAL_VALUES_EQUAL (REAL_VALUE_TYPE x, REAL_VALUE_TYPE y)**

Compares for equality the two values, x and y. If the target floating point format supports negative zeroes and/or NaNs, ‘REAL_VALUES_EQUAL (−0.0, 0.0)’ is true, and ‘REAL_VALUES_EQUAL (NaN, NaN)’ is false.

**int REAL_VALUES_LESS (REAL_VALUE_TYPE x, REAL_VALUE_TYPE y)**

Tests whether x is less than y.

**HOST_WIDE_INT REAL_VALUE_FIX (REAL_VALUE_TYPE x)**

Truncates x to a signed integer, rounding toward zero.

**unsigned HOST_WIDE_INT REAL_VALUE_UNSIGNED_FIX (REAL_VALUE_TYPE x)**

Truncates x to an unsigned integer, rounding toward zero. If x is negative, returns zero.

**REAL_VALUE_TYPE REAL_VALUE_ATOF (const char *string, enum machine_mode mode)**

Converts string into a floating point number in the target machine’s representation for mode `mode`. This routine can handle both decimal and hexadecimal floating point constants, using the syntax defined by the C language for both.

**int REAL_VALUE_NEGATIVE (REAL_VALUE_TYPE x)**

Returns 1 if x is negative (including negative zero), 0 otherwise.

**int REAL_VALUE_ISINF (REAL_VALUE_TYPE x)**

Determines whether x represents infinity (positive or negative).

**int REAL_VALUE_ISNAN (REAL_VALUE_TYPE x)**

Determines whether x represents a “NaN” (not-a-number).

**void REAL_ARITHMETIC (REAL_VALUE_TYPE output, enum tree_code code, REAL_VALUE_TYPE x, REAL_VALUE_TYPE y)**

Calculates an arithmetic operation on the two floating point values x and y, storing the result in `output` (which must be a variable).
The operation to be performed is specified by `code`. Only the following codes are supported: `PLUS_EXPR`, `MINUS_EXPR`, `MULT_EXPR`, `RDIV_EXPR`, `MAX_EXPR`, `MIN_EXPR`.

If `REAL_ARITHMETIC` is asked to evaluate division by zero and the target’s floating point format cannot represent infinity, it will call `abort`. Callers should check for this situation first, using `MODE_HAS_INFINITIES`. See Section 17.5 [Storage Layout], page 404.

```
REAL_VALUE_TYPE REAL_VALUE_NEGATE (REAL_VALUE_TYPE x)  [Macro]
Returns the negative of the floating point value x.

REAL_VALUE_TYPE REAL_VALUE_ABS (REAL_VALUE_TYPE x)  [Macro]
Returns the absolute value of x.

REAL_VALUE_TYPE REAL_VALUE_TRUNCATE (REAL_VALUE_TYPE mode, enum machine_mode x)  [Macro]
Truncates the floating point value x to fit in mode. The return value is still a full-size REAL_VALUE_TYPE, but it has an appropriate bit pattern to be output as a floating constant whose precision accords with mode mode.

void REAL_VALUE_TO_INT (HOST_WIDE_INT low, HOST_WIDE_INT high, REAL_VALUE_TYPE x)  [Macro]
Converts a floating point value x into a double-precision integer which is then stored into low and high. If the value is not integral, it is truncated.

void REAL_VALUE_FROM_INT (REAL_VALUE_TYPE x, HOST_WIDE_INT low, HOST_WIDE_INT high, enum machine_mode mode)  [Macro]
Converts a double-precision integer found in low and high, into a floating point value which is then stored into x. The value is truncated to fit in mode mode.
```

### 17.24 Mode Switching Instructions

The following macros control mode switching optimizations:

```
OPTIMIZE_MODE_SWITCHING (entity)  [Macro]
Define this macro if the port needs extra instructions inserted for mode switching in an optimizing compilation.

For an example, the SH4 can perform both single and double precision floating point operations, but to perform a single precision operation, the FPSCR PR bit has to be cleared, while for a double precision operation, this bit has to be set. Changing the PR bit requires a general purpose register as a scratch register, hence these FPSCR sets have to be inserted before reload, i.e. you can’t put this into instruction emitting or `TARGET_MACHINE_DEPENDENT_REORG`.

You can have multiple entities that are mode-switched, and select at run time which entities actually need it. `OPTIMIZE_MODE_SWITCHING` should return nonzero for any `entity` that needs mode-switching. If you define this macro, you also have to define `NUM_MODES_FOR_MODE_SWITCHING`, `MODE_NEEDED`, `MODE_PRIORITY_TO_MODE` and `EMIT_MODE_SET`. `MODE_AFTER`, `MODE_ENTRY`, and `MODE_EXIT` are optional.
```
If you define `OPTIMIZE_MODE_SWITCHING`, you have to define this as initializer for an array of integers. Each initializer element N refers to an entity that needs mode switching, and specifies the number of different modes that might need to be set for this entity. The position of the initializer in the initializer—starting counting at zero—determines the integer that is used to refer to the mode-switched entity in question. In macros that take mode arguments / yield a mode result, modes are represented as numbers 0 \ldots N – 1. N is used to specify that no mode switch is needed / supplied.

**MODE_NEEDED (entity, insn)**

`entity` is an integer specifying a mode-switched entity. If `OPTIMIZE_MODE_SWITCHING` is defined, you must define this macro to return an integer value not larger than the corresponding element in `NUM_MODES_FOR_MODE_SWITCHING`, to denote the mode that `entity` must be switched into prior to the execution of `insn`.

**MODE_AFTER (mode, insn)**

If this macro is defined, it is evaluated for every `insn` during mode switching. It determines the mode that an insn results in (if different from the incoming mode).

**MODE_ENTRY (entity)**

If this macro is defined, it is evaluated for every `entity` that needs mode switching. It should evaluate to an integer, which is a mode that `entity` is assumed to be switched to at function entry. If `MODE_ENTRY` is defined then `MODE_EXIT` must be defined.

**MODE_EXIT (entity)**

If this macro is defined, it is evaluated for every `entity` that needs mode switching. It should evaluate to an integer, which is a mode that `entity` is assumed to be switched to at function exit. If `MODE_EXIT` is defined then `MODE_ENTRY` must be defined.

**MODE_PRIORITY_TO_MODE (entity, n)**

This macro specifies the order in which modes for `entity` are processed. 0 is the highest priority, `NUM_MODES_FOR_MODE_SWITCHING[entity]` – 1 the lowest. The value of the macro should be an integer designating a mode for `entity`. For any fixed `entity`, `mode_priority_to_mode (entity, n)` shall be a bijection in 0 \ldots `num_modes_for_mode_switching[entity]` – 1.

**EMIT_MODE_SET (entity, mode, hard_regs_live)**

Generate one or more insns to set `entity` to `mode`. `hard_reg_live` is the set of hard registers live at the point where the insn(s) are to be inserted.

### 17.25 Defining target-specific uses of `__attribute__`

Target-specific attributes may be defined for functions, data and types. These are described using the following target hooks; they also need to be documented in `extend.texi`.

**const struct attribute_spec * TARGET_ATTRIBUTE_TABLE**

If defined, this target hook points to an array of `struct attribute_spec` (defined in `tree.h`) specifying the machine specific attributes for this target and some of the restrictions on the entities to which these attributes are applied and the arguments they take.
bool TARGET_ATTRIBUTE_TAKES_IDENTIFIER_P (const tree name)  [Target Hook]
If defined, this target hook is a function which returns true if the machine-specific
attribute named name expects an identifier given as its first argument to be passed on
as a plain identifier, not subjected to name lookup. If this is not defined, the default
is false for all machine-specific attributes.

int TARGET_COMP_TYPE_ATTRIBUTES (const tree type1, const tree type2)  [Target Hook]
If defined, this target hook is a function which returns zero if the attributes on type1
and type2 are incompatible, one if they are compatible, and two if they are nearly
compatible (which causes a warning to be generated). If this is not defined, machine-
specific attributes are supposed always to be compatible.

void TARGET_SET_DEFAULT_TYPE_ATTRIBUTES (tree type)  [Target Hook]
If defined, this target hook is a function which assigns default attributes to the newly
defined type.

tree TARGET_MERGE_TYPE_ATTRIBUTES (tree type1, tree type2)  [Target Hook]
Define this target hook if the merging of type attributes needs special handling. If
defined, the result is a list of the combined TYPE_ATTRIBUTES of type1 and type2. It
is assumed that comptypes has already been called and returned 1. This function
may call merge_attributes to handle machine-independent merging.

tree TARGET_MERGE_DECL_ATTRIBUTES (tree olddecl, tree newdecl)  [Target Hook]
Define this target hook if the merging of decl attributes needs special handling. If
defined, the result is a list of the combined DECL_ATTRIBUTES of olddecl and newdecl.
newdecl is a duplicate declaration of olddecl. Examples of when this is needed are
when one attribute overrides another, or when an attribute is nullified by a subsequent
definition. This function may call merge_attributes to handle machine-independent
merging.

If the only target-specific handling you require is 'dllimport' for Microsoft Win-
dows targets, you should define the macro TARGET_DLLIMPORT_DECL_ATTRIBUTES to 1.
The compiler will then define a function called merge_dllimport_decl_attributes
which can then be defined as the expansion of TARGET_MERGE_DECL_ATTRIBUTES. You
can also add handle_dll_attribute in the attribute table for your port to perform
initial processing of the 'dllimport' and 'dllexport' attributes. This is done in
'i386/cygwin.h' and 'i386/i386.c', for example.

bool TARGET_VALID_DLLIMPORT_ATTRIBUTE_P (const tree decl)  [Target Hook]
decl is a variable or function with __attribute__((dllimport)) specified. Use this
hook if the target needs to add extra validation checks to handle_dll_attribute.

TARGET_DECLSPEC  [Macro]
Define this macro to a nonzero value if you want to treat __declspec(X) as equivalent
to __attribute__((X)). By default, this behavior is enabled only for targets that define
TARGET_DLLIMPORT_DECL_ATTRIBUTES. The current implementation of __declspec
is via a built-in macro, but you should not rely on this implementation detail.
void TARGET_INSERT_ATTRIBUTES (tree node, tree *attr_ptr)   [Target Hook]
Define this target hook if you want to be able to add attributes to a decl when it
is being created. This is normally useful for back ends which wish to implement a
pragma by using the attributes which correspond to the pragma’s effect. The node
argument is the decl which is being created. The attr_ptr argument is a pointer to
the attribute list for this decl. The list itself should not be modified, since it may be
shared with other decls, but attributes may be chained on the head of the list and
*attr_ptr modified to point to the new attributes, or a copy of the list may be made
if further changes are needed.

bool TARGET_FUNCTION_ATTRIBUTE_INLINABLE_P (const tree
  fndecl)   [Target Hook]
This target hook returns true if it is ok to inline fndecl into the current function,
despite its having target-specific attributes, false otherwise. By default, if a function
has a target specific attribute attached to it, it will not be inlined.

bool TARGET_OPTION_VALID_ATTRIBUTE_P (tree fndecl, tree
  name, tree args, int flags)   [Target Hook]
This hook is called to parse the attribute(option("...")), and it allows the func-
tion to set different target machine compile time options for the current function that
might be different than the options specified on the command line. The hook should
return true if the options are valid.

  The hook should set the DECL_FUNCTION_SPECIFIC_TARGET field in the func-
tion declaration to hold a pointer to a target specific struct cl_target_option structure.

void TARGET_OPTION_SAVE (struct cl_target_option *ptr)   [Target Hook]
This hook is called to save any additional target specific information in the struct
cl_target_option structure for function specific options. See Section 8.1 [Option file
format], page 103.

void TARGET_OPTION_RESTORE (struct cl_target_option *ptr)   [Target Hook]
This hook is called to restore any additional target specific information in the struct
cl_target_option structure for function specific options.

void TARGET_OPTION_PRINT (FILE *file, int indent, struct
  cl_target_option *ptr)   [Target Hook]
This hook is called to print any additional target specific information in the struct
cl_target_option structure for function specific options.

bool TARGET_OPTION_PRAGMA_PARSE (tree args, tree pop_target)   [Target Hook]
This target hook parses the options for #pragma GCC option to set the machine spe-
cific options for functions that occur later in the input stream. The options should
be the same as handled by the TARGET_OPTION_VALID_ATTRIBUTE_P hook.

void TARGET_OPTION_OVERRIDE (void)   [Target Hook]
Sometimes certain combinations of command options do not make sense on a particular
target machine. You can override the hook TARGET_OPTION_OVERRIDE to take
account of this. This hooks is called once just after all the command options have
been parsed.
Don’t use this hook to turn on various extra optimizations for ‘-O’. That is what TARGET_OPTION_OPTIMIZATION is for.

If you need to do something whenever the optimization level is changed via the optimize attribute or pragma, see TARGET_OVERRIDE_OPTIONS_AFTER_CHANGE

bool TARGET_CAN_INLINE_P (tree caller, tree callee) [Target Hook]
This target hook returns false if the caller function cannot inline callee, based on target specific information. By default, inlining is not allowed if the callee function has function specific target options and the caller does not use the same options.

17.26 Emulating TLS

For targets whose psABI does not provide Thread Local Storage via specific relocations and instruction sequences, an emulation layer is used. A set of target hooks allows this emulation layer to be configured for the requirements of a particular target. For instance the psABI may in fact specify TLS support in terms of an emulation layer.

The emulation layer works by creating a control object for every TLS object. To access the TLS object, a lookup function is provided which, when given the address of the control object, will return the address of the current thread’s instance of the TLS object.

const char * TARGET_EMUTLS_GET_ADDRESS [Target Hook]
Contains the name of the helper function that uses a TLS control object to locate a TLS instance. The default causes libgcc’s emulated TLS helper function to be used.

const char * TARGET_EMUTLS_REGISTER_COMMON [Target Hook]
Contains the name of the helper function that should be used at program startup to register TLS objects that are implicitly initialized to zero. If this is NULL, all TLS objects will have explicit initializers. The default causes libgcc’s emulated TLS registration function to be used.

const char * TARGET_EMUTLS_VAR_SECTION [Target Hook]
Contains the name of the section in which TLS control variables should be placed. The default of NULL allows these to be placed in any section.

const char * TARGET_EMUTLS_TMPL_SECTION [Target Hook]
Contains the name of the section in which TLS initializers should be placed. The default of NULL allows these to be placed in any section.

const char * TARGET_EMUTLS_VAR_PREFIX [Target Hook]
Contains the prefix to be prepended to TLS control variable names. The default of NULL uses a target-specific prefix.

const char * TARGET_EMUTLS_TMPL_PREFIX [Target Hook]
Contains the prefix to be prepended to TLS initializer objects. The default of NULL uses a target-specific prefix.

tree TARGET_EMUTLS_VAR_FIELDS (tree type, tree *name) [Target Hook]
Specifies a function that generates the FIELD_DECLS for a TLS control object type. type is the RECORD_TYPE the fields are for and name should be filled with the structure tag, if the default of __emutls_object is unsuitable. The default creates a type suitable for libgcc’s emulated TLS function.
tree TARGET_EMUTLS_VAR_INIT (tree var, tree decl, tree tmpl_addr)  
[Target Hook]
Specifies a function that generates the CONSTRUCTOR to initialize a TLS control object. var is the TLS control object, decl is the TLS object and tmpl_addr is the address of the initializer. The default initializes libgcc’s emulated TLS control object.

bool TARGET_EMUTLS_VAR_ALIGN_FIXED  
[Target Hook]
Specifies whether the alignment of TLS control variable objects is fixed and should not be increased as some backends may do to optimize single objects. The default is false.

bool TARGET_EMUTLS_DEBUG_FORM_TLS_ADDRESS  
[Target Hook]
Specifies whether a DWARF DW_OP_form_tls_address location descriptor may be used to describe emulated TLS control objects.

17.27 Defining coprocessor specifics for MIPS targets.

The MIPS specification allows MIPS implementations to have as many as 4 coprocessors, each with as many as 32 private registers. GCC supports accessing these registers and transferring values between the registers and memory using asm-ized variables. For example:

```
register unsigned int cp0count asm ("c0r1");
unsigned int d;

d = cp0count + 3;
```

(“c0r1” is the default name of register 1 in coprocessor 0; alternate names may be added as described below, or the default names may be overridden entirely in SUBTARGET_CONDITIONAL_REGISTER_USAGE.)

Coprocessor registers are assumed to be epilogue-used; sets to them will be preserved even if it does not appear that the register is used again later in the function.

Another note: according to the MIPS spec, coprocessor 1 (if present) is the FPU. One accesses COP1 registers through standard mips floating-point support; they are not included in this mechanism.

There is one macro used in defining the MIPS coprocessor interface which you may want to override in subtargets; it is described below.

**ALL_COP_ADDITIONAL_REGISTER_NAMES**  
[Macro]
A comma-separated list (with leading comma) of pairs describing the alternate names of coprocessor registers. The format of each entry should be

```
{ alternatename, register_number}
```

Default: empty.

17.28 Parameters for Precompiled Header Validity Checking

void * TARGET_GET_PCH_VALIDITY (size_t *sz)  
[Target Hook]
This hook returns a pointer to the data needed by TARGET_PCH_VALID_P and sets ‘*sz’ to the size of the data in bytes.
const char * TARGET_PCH_VALID_P (const void *data, size_t sz)  [Target Hook]
This hook checks whether the options used to create a PCH file are compatible with the current settings. It returns NULL if so and a suitable error message if not. Error messages will be presented to the user and must be localized using '_(msg)'.
data is the data that was returned by TARGET_GET_PCH_VALIDITY when the PCH file was created and sz is the size of that data in bytes. It’s safe to assume that the data was created by the same version of the compiler, so no format checking is needed.
The default definition of default_pch_valid_p should be suitable for most targets.

const char * TARGET_CHECK_PCH_TARGET_FLAGS (int pch_flags)  [Target Hook]
If this hook is nonnull, the default implementation of TARGET_PCH_VALID_P will use it to check for compatible values of target_flags. pch_flags specifies the value that target_flags had when the PCH file was created. The return value is the same as for TARGET_PCH_VALID_P.

### 17.29 C++ ABI parameters

tree TARGET_CXX_GUARD_TYPE (void)  [Target Hook]
Define this hook to override the integer type used for guard variables. These are used to implement one-time construction of static objects. The default is long long integer type node.

bool TARGET_CXX_GUARD_MASK_BIT (void)  [Target Hook]
This hook determines how guard variables are used. It should return false (the default) if the first byte should be used. A return value of true indicates that only the least significant bit should be used.

tree TARGET_CXX_GET_COOKIE_SIZE (tree type)  [Target Hook]
This hook returns the size of the cookie to use when allocating an array whose elements have the indicated type. Assumes that it is already known that a cookie is needed. The default is max(sizeof (size_t), alignof(type)), as defined in section 2.7 of the IA64/Generic C++ ABI.

bool TARGET_CXX_COOKIE_HAS_SIZE (void)  [Target Hook]
This hook should return true if the element size should be stored in array cookies. The default is to return false.

int TARGET_CXX_IMPORT_EXPORT_CLASS (tree type, int import_export)  [Target Hook]
If defined by a backend this hook allows the decision made to export class type to be overruled. Upon entry import_export will contain 1 if the class is going to be exported, -1 if it is going to be imported and 0 otherwise. This function should return the modified value and perform any other actions necessary to support the backend’s targeted operating system.

bool TARGET_CXX_CDTOR_RETURNS_THIS (void)  [Target Hook]
This hook should return true if constructors and destructors return the address of the object created/destroyed. The default is to return false.
bool TARGET_CXX_KEY_METHOD_MAY_BE_INLINE (void) [Target Hook]
This hook returns true if the key method for a class (i.e., the method which, if defined in the current translation unit, causes the virtual table to be emitted) may be an inline function. Under the standard Itanium C++ ABI the key method may be an inline function so long as the function is not declared inline in the class definition. Under some variants of the ABI, an inline function can never be the key method. The default is to return true.

void TARGET_CXX_DETERMINE_CLASS_DATA_VISIBILITY (tree decl) [Target Hook]
dcl is a virtual table, virtual table table, typeinfo object, or other similar implicit class data object that will be emitted with external linkage in this translation unit. No ELF visibility has been explicitly specified. If the target needs to specify a visibility other than that of the containing class, use this hook to set DECL_VISIBILITY and DECL_VISIBILITY_Specified.

bool TARGET_CXX_CLASS_DATA_ALWAYS_COMDAT (void) [Target Hook]
This hook returns true (the default) if virtual tables and other similar implicit class data objects are always COMDAT if they have external linkage. If this hook returns false, then class data for classes whose virtual table will be emitted in only one translation unit will not be COMDAT.

bool TARGET_CXX_LIBRARY_RTTI_COMDAT (void) [Target Hook]
This hook returns true (the default) if the RTTI information for the basic types which is defined in the C++ runtime should always be COMDAT, false if it should not be COMDAT.

bool TARGET_CXX_USE_AEABI_AEXIT (void) [Target Hook]
This hook returns true if __aeabi_atexit (as defined by the ARM EABI) should be used to register static destructors when -fuse-cxa-atexit is in effect. The default is to return false to use __cxa_atexit.

bool TARGET_CXX_USE_AEXIT_FOR_CXA_AEXIT (void) [Target Hook]
This hook returns true if the target atexit function can be used in the same manner as __cxa_atexit to register C++ static destructors. This requires that atexit-registered functions in shared libraries are run in the correct order when the libraries are unloaded. The default is to return false.

void TARGET_CXX_ADJUST_CLASS_AT_DEFINITION (tree type) [Target Hook]
type is a C++ class (i.e., RECORD_TYPE or UNION_TYPE) that has just been defined. Use this hook to make adjustments to the class (e.g., tweak visibility or perform any other required target modifications).

17.30 Adding support for named address spaces
The draft technical report of the ISO/IEC JTC1 S22 WG14 N1275 standards committee, *Programming Languages - C - Extensions to support embedded processors*, specifies a syntax for embedded processors to specify alternate address spaces. You can configure a GCC port to support section 5.1 of the draft report to add support for address spaces other than
the default address space. These address spaces are new keywords that are similar to the volatile and const type attributes.

Pointers to named address spaces can have a different size than pointers to the generic address space.

For example, the SPU port uses the __ea address space to refer to memory in the host processor, rather than memory local to the SPU processor. Access to memory in the __ea address space involves issuing DMA operations to move data between the host processor and the local processor memory address space. Pointers in the __ea address space are either 32 bits or 64 bits based on the ‘-mea32’ or ‘-mea64’ switches (native SPU pointers are always 32 bits).

Internally, address spaces are represented as a small integer in the range 0 to 15 with address space 0 being reserved for the generic address space.

To register a named address space qualifier keyword with the C front end, the target may call the c_register_addr_space routine. For example, the SPU port uses the following to declare __ea as the keyword for named address space #1:

```
#define ADDR_SPACE_EA 1
  c_register_addr_space("__ea", ADDR_SPACE_EA);
```

```
enum machine_mode TARGET_ADDR_SPACE_POINTER_MODE [Target Hook]
  (addr_space_t address_space)
  Define this to return the machine mode to use for pointers to address_space if the target supports named address spaces. The default version of this hook returns ptr_mode for the generic address space only.

enum machine_mode TARGET_ADDR_SPACE_ADDRESS_MODE [Target Hook]
  (addr_space_t address_space)
  Define this to return the machine mode to use for addresses in address_space if the target supports named address spaces. The default version of this hook returns Pmode for the generic address space only.

bool TARGET_ADDR_SPACE_VALID_POINTER_MODE (enum
  machine_mode mode, addr_space_t as) [Target Hook]
  Define this to return nonzero if the port can handle pointers with machine mode mode to address space as. This target hook is the same as the TARGET_VALID_POINTER_MODE target hook, except that it includes explicit named address space support. The default version of this hook returns true for the modes returned by either the TARGET_ADDR_SPACE_POINTER_MODE or TARGET_ADDR_SPACE_ADDRESS_MODE target hooks for the given address space.

bool TARGET_ADDR_SPACE_LEGITIMATE_ADDRESS_P (enum
  machine_mode mode, rtx exp, bool strict, addr_space_t as) [Target Hook]
  Define this to return true if exp is a valid address for mode mode in the named address space as. The strict parameter says whether strict addressing is in effect after reload has finished. This target hook is the same as the TARGET_LEGITIMATE_ADDRESS_P target hook, except that it includes explicit named address space support.
rtx TARGET_ADDR_SPACE_LEGITIMIZE_ADDRESS (rtx x, rtx oldx,      [Target Hook]
    enum machine_mode mode, addr_space_t as)
    Define this to modify an invalid address x to be a valid address with mode mode in the
named address space as. This target hook is the same as the TARGET_LEGITIMIZE_ADDRESS target hook, except that it includes explicit named address space support.

bool TARGET_ADDR_SPACE_SUBSET_P (addr_space_t superset,
    addr_space_t subset)          [Target Hook]
    Define this to return whether the subset named address space is contained within the
superset named address space. Pointers to a named address space that is a subset of another named address space will be converted automatically without a cast if used together in arithmetic operations. Pointers to a superset address space can be converted to pointers to a subset address space via explicit casts.

rtx TARGET_ADDR_SPACE_CONVERT (rtx op, tree from_type, tree
    to_type)                       [Target Hook]
    Define this to convert the pointer expression represented by the RTL op with type
from_type that points to a named address space to a new pointer expression with type
to_type that points to a different named address space. When this hook it called, it is
guaranteed that one of the two address spaces is a subset of the other, as determined
by the TARGET_ADDR_SPACE_SUBSET_P target hook.

17.31 Miscellaneous Parameters

Here are several miscellaneous parameters.

HAS_LONG_COND_BRANCH            [Macro]
    Define this boolean macro to indicate whether or not your architecture has conditional
branches that can span all of memory. It is used in conjunction with an optimization
that partitions hot and cold basic blocks into separate sections of the executable. If
this macro is set to false, gcc will convert any conditional branches that attempt to
cross between sections into unconditional branches or indirect jumps.

HAS_LONG_UNCOND_BRANCH          [Macro]
    Define this boolean macro to indicate whether or not your architecture has unconditional
branches that can span all of memory. It is used in conjunction with an
optimization that partitions hot and cold basic blocks into separate sections of the
executable. If this macro is set to false, gcc will convert any unconditional branches
that attempt to cross between sections into indirect jumps.

CASE_VECTOR_MODE                [Macro]
    An alias for a machine mode name. This is the machine mode that elements of a
jump-table should have.

CASE_VECTOR_SHORTEN_MODE (min_offset, max_offset, body)      [Macro]
    Optional: return the preferred mode for an addr_diff_vec when the minimum and
maximum offset are known. If you define this, it enables extra code in branch shortening
to deal with addr_diff_vec. To make this work, you also have to define INSN_ALIGN and make the alignment for addr_diff_vec explicit. The body argument is provided so that the offset_unsigned and scale flags can be updated.
**CASE_VECTOR_PC_RELATIVE**

Define this macro to be a C expression to indicate when jump-tables should contain relative addresses. You need not define this macro if jump-tables never contain relative addresses, or jump-tables should contain relative addresses only when ‘-fPIC’ or ‘-fPIC’ is in effect.

**unsigned int TARGET_CASE_VALUES_THRESHOLD (void)**

This function return the smallest number of different values for which it is best to use a jump-table instead of a tree of conditional branches. The default is four for machines with a casesi instruction and five otherwise. This is best for most machines.

**CASE_USE_BIT_TESTS**

Define this macro to be a C expression to indicate whether C switch statements may be implemented by a sequence of bit tests. This is advantageous on processors that can efficiently implement left shift of 1 by the number of bits held in a register, but inappropriate on targets that would require a loop. By default, this macro returns true if the target defines an ashlsi3 pattern, and false otherwise.

**WORD_REGISTER_OPERATIONS**

Define this macro if operations between registers with integral mode smaller than a word are always performed on the entire register. Most RISC machines have this property and most CISC machines do not.

**LOAD_EXTEND_OP (mem_mode)**

Define this macro to be a C expression indicating when inssns that read memory in mem_mode, an integral mode narrower than a word, set the bits outside of mem_mode to be either the sign-extension or the zero-extension of the data read. Return SIGN_EXTEND for values of mem_mode for which the insn sign-extends, ZERO_EXTEND for which it zero-extends, and UNKNOWN for other modes.

This macro is not called with mem_mode non-integral or with a width greater than or equal to BITS_PER_WORD, so you may return any value in this case. Do not define this macro if it would always return UNKNOWN. On machines where this macro is defined, you will normally define it as the constant SIGN_EXTEND or ZERO_EXTEND.

You may return a non-UNKNOWN value even if for some hard registers the sign extension is not performed, if for the REGNO_REG_CLASS of these hard registers CANNOT_CHANGE_MODE_CLASS returns nonzero when the from mode is mem_mode and the to mode is any integral mode larger than this but not larger than word_mode.

You must return UNKNOWN if for some hard registers that allow this mode, CANNOT_CHANGE_MODE_CLASS says that they cannot change to word_mode, but that they can change to another integral mode that is larger then mem_mode but still smaller than word_mode.

**SHORT_IMMEDIATES_SIGN_EXTEND**

Define this macro if loading short immediate values into registers sign extends.

**FIXUNS_TRUNC_LIKE_FIX_TRUNC**

Define this macro if the same instructions that convert a floating point number to a signed fixed point number also convert validly to an unsigned one.
unsigned int TARGET_MIN_DIVISIONS_FOR_RECIP_MUL (enum machine_mode mode) [Target Hook]

When ‘-ffast-math’ is in effect, GCC tries to optimize divisions by the same divisor, by turning them into multiplications by the reciprocal. This target hook specifies the minimum number of divisions that should be there for GCC to perform the optimization for a variable of mode mode. The default implementation returns 3 if the machine has an instruction for the division, and 2 if it does not.

MOVE_MAX [Macro]

The maximum number of bytes that a single instruction can move quickly between memory and registers or between two memory locations.

MAX_MOVE_MAX [Macro]

The maximum number of bytes that a single instruction can move quickly between memory and registers or between two memory locations. If this is undefined, the default is MOVE_MAX. Otherwise, it is the constant value that is the largest value that MOVE_MAX can have at run-time.

SHIFT_COUNT_TRUNCATED [Macro]

A C expression that is nonzero if on this machine the number of bits actually used for the count of a shift operation is equal to the number of bits needed to represent the size of the object being shifted. When this macro is nonzero, the compiler will assume that it is safe to omit a sign-extend, zero-extend, and certain bitwise ‘and’ instructions that truncates the count of a shift operation. On machines that have instructions that act on bit-fields at variable positions, which may include ‘bit test’ instructions, a nonzero SHIFT_COUNT_TRUNCATED also enables deletion of truncations of the values that serve as arguments to bit-field instructions.

If both types of instructions truncate the count (for shifts) and position (for bit-field operations), or if no variable-position bit-field instructions exist, you should define this macro.

However, on some machines, such as the 80386 and the 680x0, truncation only applies to shift operations and not the (real or pretended) bit-field operations. Define SHIFT_COUNT_TRUNCATED to be zero on such machines. Instead, add patterns to the ‘md’ file that include the implied truncation of the shift instructions.

You need not define this macro if it would always have the value of zero.

unsigned HOST_WIDE_INT TARGET_SHIFT_TRUNCATION_MASK (enum machine_mode mode) [Target Hook]

This function describes how the standard shift patterns for mode deal with shifts by negative amounts or by more than the width of the mode. See [shift patterns], page 339.

On many machines, the shift patterns will apply a mask m to the shift count, meaning that a fixed-width shift of x by y is equivalent to an arbitrary-width shift of x by y & m. If this is true for mode mode, the function should return m, otherwise it should return 0. A return value of 0 indicates that no particular behavior is guaranteed.

Note that, unlike SHIFT_COUNT_TRUNCATED, this function does not apply to general shift rtxes; it applies only to instructions that are generated by the named shift patterns.
The default implementation of this function returns \( \text{GET\_MODE\_BITSIZE}(\text{mode}) - 1 \) if \( \text{SHIFT\_COUNT\_TRUNCATED} \) and 0 otherwise. This definition is always safe, but if \( \text{SHIFT\_COUNT\_TRUNCATED} \) is false, and some shift patterns nevertheless truncate the shift count, you may get better code by overriding it.

**TRULY\_NOOP\_TRUNCATION** \((\text{outprec}, \text{inprec})\)  
**[Macro]**  
A C expression which is nonzero if on this machine it is safe to “convert” an integer of \( \text{inprec} \) bits to one of \( \text{outprec} \) bits (where \( \text{outprec} \) is smaller than \( \text{inprec} \)) by merely operating on it as if it had only \( \text{outprec} \) bits. 

On many machines, this expression can be 1.  
When **TRULY\_NOOP\_TRUNCATION** returns 1 for a pair of sizes for modes for which \( \text{MODES\_TIEABLE\_P} \) is 0, suboptimal code can result. If this is the case, making **TRULY\_NOOP\_TRUNCATION** return 0 in such cases may improve things.

**int TARGET\_MODE\_REP\_EXTENDED** \((\text{enum machine\_mode, enum mode}, \text{enum machine\_mode rep\_mode})\)  
**[Target Hook]**  
The representation of an integral mode can be such that the values are always extended to a wider integral mode. Return \( \text{SIGN\_EXTEND} \) if values of \( \text{mode} \) are represented in sign-extended form to \( \text{rep\_mode} \). Return \( \text{UNKNOWN} \) otherwise. (Currently, none of the targets use zero-extended representation this way so unlike \( \text{LOAD\_EXTEND\_OP} \), \( \text{TARGTHET\_MODE\_REP\_EXTENDED} \) is expected to return either \( \text{SIGN\_EXTEND} \) or \( \text{UNKNOWN} \). Also no target extends \( \text{mode} \) to \( \text{rep\_mode} \) so that \( \text{rep\_mode} \) is not the next widest integral mode and currently we take advantage of this fact.)  
Similarly to \( \text{LOAD\_EXTEND\_OP} \) you may return a non-\( \text{UNKNOWN} \) value even if the extension is not performed on certain hard registers as long as for the \( \text{REGNO\_REG\_CLASS} \) of these hard registers \( \text{CANNOT\_CHANGE\_MODE\_CLASS} \) returns nonzero.  
Note that \( \text{TARGET\_MODE\_REP\_EXTENDED} \) and \( \text{LOAD\_EXTEND\_OP} \) describe two related properties. If you define \( \text{TARGET\_MODE\_REP\_EXTENDED}(\text{mode, word\_mode}) \) you probably also want to define \( \text{LOAD\_EXTEND\_OP}(\text{mode}) \) to return the same type of extension.  
In order to enforce the representation of \( \text{mode} \), **TRULY\_NOOP\_TRUNCATION** should return false when truncating to \( \text{mode} \).

**STORE\_FLAG\_VALUE**  
**[Macro]**  
A C expression describing the value returned by a comparison operator with an integral mode and stored by a store-flag instruction (‘\text{cstore\_mode4}’) when the condition is true. This description must apply to all the ‘\text{cstore\_mode4}’ patterns and all the comparison operators whose results have a \( \text{MODE\_INT} \) mode.  
A value of 1 or \(-1\) means that the instruction implementing the comparison operator returns exactly 1 or \(-1\) when the comparison is true and 0 when the comparison is false. Otherwise, the value indicates which bits of the result are guaranteed to be 1 when the comparison is true. This value is interpreted in the mode of the comparison operation, which is given by the mode of the first operand in the ‘\text{cstore\_mode4}’ pattern. Either the low bit or the sign bit of \( \text{STORE\_FLAG\_VALUE} \) be on. Presently, only those bits are used by the compiler.  
If \( \text{STORE\_FLAG\_VALUE} \) is neither 1 or \(-1\), the compiler will generate code that depends only on the specified bits. It can also replace comparison operators with equivalent
operations if they cause the required bits to be set, even if the remaining bits are undefined. For example, on a machine whose comparison operators return an SImode value and where STORE_FLAG_VALUE is defined as ‘0x80000000’, saying that just the sign bit is relevant, the expression

\[(\text{ne:SI } (\text{and:SI } x \ (\text{const_int power-of-2}))) \ (\text{const_int 0})\]

can be converted to

\[(\text{ashift:SI } x \ (\text{const_int n}))\]

where \(n\) is the appropriate shift count to move the bit being tested into the sign bit.

There is no way to describe a machine that always sets the low-order bit for a true value, but does not guarantee the value of any other bits, but we do not know of any machine that has such an instruction. If you are trying to port GCC to such a machine, include an instruction to perform a logical-and of the result with 1 in the pattern for the comparison operators and let us know at gcc@gcc.gnu.org.

Often, a machine will have multiple instructions that obtain a value from a comparison (or the condition codes). Here are rules to guide the choice of value for STORE_FLAG_VALUE, and hence the instructions to be used:

- Use the shortest sequence that yields a valid definition for STORE_FLAG_VALUE. It is more efficient for the compiler to “normalize” the value (convert it to, e.g., 1 or 0) than for the comparison operators to do so because there may be opportunities to combine the normalization with other operations.
- For equal-length sequences, use a value of 1 or \(-1\), with \(-1\) being slightly preferred on machines with expensive jumps and 1 preferred on other machines.
- As a second choice, choose a value of ‘0x80000001’ if instructions exist that set both the sign and low-order bits but do not define the others.
- Otherwise, use a value of ‘0x80000000’.

Many machines can produce both the value chosen for STORE_FLAG_VALUE and its negation in the same number of instructions. On those machines, you should also define a pattern for those cases, e.g., one matching

\[(\text{set A } (\text{neg:m } (\text{ne:m B C})))\]

Some machines can also perform and or plus operations on condition code values with less instructions than the corresponding ‘cstoremode4’ insn followed by and or plus. On those machines, define the appropriate patterns. Use the names incsc and decsc, respectively, for the patterns which perform plus or minus operations on condition code values. See ‘rs6000.md’ for some examples. The GNU Superoptimizer can be used to find such instruction sequences on other machines.

If this macro is not defined, the default value, 1, is used. You need not define STORE_FLAG_VALUE if the machine has no store-flag instructions, or if the value generated by these instructions is 1.

**FLOAT_STORE_FLAG_VALUE (mode)** [Macro]

A C expression that gives a nonzero REAL_VALUE_TYPE value that is returned when comparison operators with floating-point results are true. Define this macro on machines that have comparison operations that return floating-point values. If there are no such operations, do not define this macro.
VECTOR_STORE_FLAG_VALUE (mode) [Macro]
A C expression that gives a rtx representing the nonzero true element for vector comparisons. The returned rtx should be valid for the inner mode of mode which is guaranteed to be a vector mode. Define this macro on machines that have vector comparison operations that return a vector result. If there are no such operations, do not define this macro. Typically, this macro is defined as const1_rtx or constm1_rtx. This macro may return NULL_RTX to prevent the compiler optimizing such vector comparison operations for the given mode.

CLZ_DEFINED_VALUE_AT_ZERO (mode, value) [Macro]
CTZ_DEFINED_VALUE_AT_ZERO (mode, value) [Macro]
A C expression that indicates whether the architecture defines a value for clz or ctz with a zero operand. A result of 0 indicates the value is undefined. If the value is defined for only the RTL expression, the macro should evaluate to 1; if the value applies also to the corresponding optab entry (which is normally the case if it expands directly into the corresponding RTL), then the macro should evaluate to 2. In the cases where the value is defined, value should be set to this value.

If this macro is not defined, the value of clz or ctz at zero is assumed to be undefined. This macro must be defined if the target’s expansion for ffs relies on a particular value to get correct results. Otherwise it is not necessary, though it may be used to optimize some corner cases, and to provide a default expansion for the ffs optab.

Note that regardless of this macro the “definedness” of clz and ctz at zero do not extend to the builtin functions visible to the user. Thus one may be free to adjust the value at will to match the target expansion of these operations without fear of breaking the API.

Pmode [Macro]
An alias for the machine mode for pointers. On most machines, define this to be the integer mode corresponding to the width of a hardware pointer; SImode on 32-bit machine or DImode on 64-bit machines. On some machines you must define this to be one of the partial integer modes, such as PSImode.

The width of Pmode must be at least as large as the value of POINTER_SIZE. If it is not equal, you must define the macro POINTERS_EXTEND_UNSIGNED to specify how pointers are extended to Pmode.

FUNCTION_MODE [Macro]
An alias for the machine mode used for memory references to functions being called, in call RTL expressions. On most CISC machines, where an instruction can begin at any byte address, this should be QImode. On most RISC machines, where all instructions have fixed size and alignment, this should be a mode with the same size and alignment as the machine instruction words - typically SImode or HImode.

STDC_0_IN_SYSTEM_HEADERS [Macro]
In normal operation, the preprocessor expands __STDC__ to the constant 1, to signify that GCC conforms to ISO Standard C. On some hosts, like Solaris, the system compiler uses a different convention, where __STDC__ is normally 0, but is 1 if the user specifies strict conformance to the C Standard.
Defining \texttt{STDC\_0\_IN\_SYSTEM\_HEADERS} makes GNU CPP follows the host convention when processing system header files, but when processing user files \texttt{__STDC__} will always expand to 1.

\textbf{NO\_IMP\_EXTERN\_C} \hspace{1cm} \textbf{[Macro]}
Define this macro if the system header files support C++ as well as C. This macro inhibits the usual method of using system header files in C++, which is to pretend that the file’s contents are enclosed in ‘\texttt{extern \"C\" \{\ldots\}’.

\textbf{REGISTER\_TARGET\_PRAGMAS ()} \hspace{1cm} \textbf{[Macro]}
Define this macro if you want to implement any target-specific pragmas. If defined, it is a C expression which makes a series of calls to \texttt{c\_register\_pragma} or \texttt{c\_register\_pragma\_with\_expansion} for each pragma. The macro may also do any setup required for the pragmas.

The primary reason to define this macro is to provide compatibility with other compilers for the same target. In general, we discourage definition of target-specific pragmas for GCC.

If the pragma can be implemented by attributes then you should consider defining the target hook ‘\texttt{T\_INSERT\_ATTRIBUTES}’ as well.

Preprocessor macros that appear on pragma lines are not expanded. All ‘\texttt{#pragma}’ directives that do not match any registered pragma are silently ignored, unless the user specifies ‘\texttt{-Wunknown\_ pragmas}’.

\begin{verbatim}
void c_registerPragma (const char *space, const char *name, void (*callback) (struct cpp_reader *))
void c_registerPragmaWithExpansion (const char *space, const char *name, void (*callback) (struct cpp_reader *))
\end{verbatim}

Each call to \texttt{c\_register\_pragma} or \texttt{c\_register\_pragma\_with\_expansion} establishes one pragma. The \texttt{callback} routine will be called when the preprocessor encounters a pragma of the form

\begin{verbatim}
#pragma [space] name ...
\end{verbatim}

\texttt{space} is the case-sensitive namespace of the pragma, or \texttt{NULL} to put the pragma in the global namespace. The callback routine receives \texttt{pfile} as its first argument, which can be passed on to \texttt{cpplib}’s functions if necessary. You can lex tokens after the \texttt{name} by calling \texttt{pragma\_lex}. Tokens that are not read by the callback will be silently ignored. The end of the line is indicated by a token of type \texttt{CPP\_EOF}. Macro expansion occurs on the arguments of pragmas registered with \texttt{c\_register\_pragma\_with\_expansion} but not on the arguments of pragmas registered with \texttt{c\_register\_pragma}.

Note that the use of \texttt{pragma\_lex} is specific to the C and C++ compilers. It will not work in the Java or Fortran compilers, or any other language compilers for that matter. Thus if \texttt{pragma\_lex} is going to be called from target-specific code, it must only be done so when building the C and C++ compilers. This can be done by defining the variables \texttt{c\_target\_objs} and \texttt{cxx\_target\_objs} in the target entry in the ‘\texttt{config\_gcc}’ file. These variables should name the target-specific, language-specific object file which contains the code that uses \texttt{pragma\_lex}. Note it will also be necessary to add a rule to the makefile fragment pointed to by \texttt{tmake\_file} that shows how to build this object file.
Chapter 17: Target Description Macros and Functions

HANDLE_PRAGMA_PACK_WITH_EXPANSION [Macro]
Define this macro if macros should be expanded in the arguments of ‘#pragma pack’.

bool TARGET_HANDLE_PRAGMA_EXTERN_PREFIX [Target Hook]
True if #pragma extern_prefix is to be supported.

TARGET_DEFAULT_PACK_STRUCT [Macro]
If your target requires a structure packing default other than 0 (meaning the machine
default), define this macro to the necessary value (in bytes). This must be a value
that would also be valid to use with ‘#pragma pack’ (that is, a small power of two).

DOLLARS_IN_IDENTIFIERS [Macro]
Define this macro to control use of the character ‘$’ in identifier names for the C
family of languages. 0 means ‘$’ is not allowed by default; 1 means it is allowed. 1 is
the default; there is no need to define this macro in that case.

NO_DOLLAR_IN_LABEL [Macro]
Define this macro if the assembler does not accept the character ‘$’ in label names.
By default constructors and destructors in G++ have ‘$’ in the identifiers. If this
macro is defined, ‘.’ is used instead.

NO_DOT_IN_LABEL [Macro]
Define this macro if the assembler does not accept the character ‘.’ in label names.
By default constructors and destructors in G++ have names that use ‘.’. If this macro
is defined, these names are rewritten to avoid ‘.’.

INSN_SETS_ARE_DELAYED (insn) [Macro]
Define this macro as a C expression that is nonzero if it is safe for the delay slot
scheduler to place instructions in the delay slot of insn, even if they appear to use
a resource set or clobbered in insn. insn is always a jump_insn or an insn; GCC
knows that every call_insn has this behavior. On machines where some insn or
jump_insn is really a function call and hence has this behavior, you should define
this macro.
You need not define this macro if it would always return zero.

INSN_REFERENCES_ARE_DELAYED (insn) [Macro]
Define this macro as a C expression that is nonzero if it is safe for the delay slot
scheduler to place instructions in the delay slot of insn, even if they appear to set or
clobber a resource referenced in insn. insn is always a jump_insn or an insn. On
machines where some insn or jump_insn is really a function call and its operands are
registers whose use is actually in the subroutine it calls, you should define this macro.
Doing so allows the delay slot scheduler to move instructions which copy arguments
into the argument registers into the delay slot of insn.
You need not define this macro if it would always return zero.

MULTIPLE_SYMBOL_SPACES [Macro]
Define this macro as a C expression that is nonzero if, in some cases, global symbols
from one translation unit may not be bound to undefined symbols in another transla-
tion unit without user intervention. For instance, under Microsoft Windows symbols
must be explicitly imported from shared libraries (DLLs).
You need not define this macro if it would always evaluate to zero.

**tree** **TARGET_MD_ASM_CLOBBERS** *(tree outputs, tree inputs, tree clobbers)*

[Target Hook]

This target hook should add to clobbers STRING_CST trees for any hard regs the port wishes to automatically clobber for an asm. It should return the result of the last `tree_cons` used to add a clobber. The `outputs`, `inputs` and `clobber` lists are the corresponding parameters to the asm and may be inspected to avoid clobbering a register that is an input or output of the asm. You can use `tree_overlaps_hard_reg_set`, declared in `tree.h`, to test for overlap with regards to asm-declared registers.

**MATH_LIBRARY**

[Macro]

Define this macro as a C string constant for the linker argument to link in the system math library, minus the initial `"-l"`, or `""` if the target does not have a separate math library.

You need only define this macro if the default of `"m"` is wrong.

**LIBRARY_PATH_ENV**

[Macro]

Define this macro as a C string constant for the environment variable that specifies where the linker should look for libraries.

You need only define this macro if the default of `"LIBRARY_PATH"` is wrong.

**TARGET_POSIX_IO**

[Macro]

Define this macro if the target supports the following POSIX file functions, access, mkdir and file locking with fcntl / F_SETLKW. Defining `TARGET_POSIX_IO` will enable the test coverage code to use file locking when exiting a program, which avoids race conditions if the program has forked. It will also create directories at run-time for cross-profiling.

**MAX_CONDITIONAL_EXECUTE**

[Macro]

A C expression for the maximum number of instructions to execute via conditional execution instructions instead of a branch. A value of `BRANCH_COST+1` is the default if the machine does not use cc0, and 1 if it does use cc0.

**IFCVT_MODIFY_TESTS** *(ce_info, true_expr, false_expr)*

[Macro]

Used if the target needs to perform machine-dependent modifications on the conditionals used for turning basic blocks into conditionally executed code. `ce_info` points to a data structure, `struct ce_if_block`, which contains information about the currently processed blocks. `true_expr` and `false_expr` are the tests that are used for converting the then-block and the else-block, respectively. Set either `true_expr` or `false_expr` to a null pointer if the tests cannot be converted.

**IFCVT_MODIFY_MULTIPLE_TESTS** *(ce_info, bb, true_expr, false_expr)*

[Macro]

Like `IFCVT_MODIFY_TESTS`, but used when converting more complicated if-statements into conditions combined by `and` and `or` operations. `bb` contains the basic block that contains the test that is currently being processed and about to be turned into a condition.
IFCVT_MODIFY_INSN (ce_info, pattern, insn) [Macro]
A C expression to modify the PATTERN of an INSN that is to be converted to conditional execution format. ce_info points to a data structure, struct ce_if_block, which contains information about the currently processed blocks.

IFCVT_MODIFY_FINAL (ce_info) [Macro]
A C expression to perform any final machine dependent modifications in converting code to conditional execution. The involved basic blocks can be found in the struct ce_if_block structure that is pointed to by ce_info.

IFCVT_MODIFY_CANCEL (ce_info) [Macro]
A C expression to cancel any machine dependent modifications in converting code to conditional execution. The involved basic blocks can be found in the struct ce_if_block structure that is pointed to by ce_info.

IFCVT_INIT_EXTRA_FIELDS (ce_info) [Macro]
A C expression to initialize any extra fields in a struct ce_if_block structure, which are defined by the IFCVT_EXTRA_FIELDS macro.

IFCVT_EXTRA_FIELDS [Macro]
If defined, it should expand to a set of field declarations that will be added to the struct ce_if_block structure. These should be initialized by the IFCVT_INIT_EXTRA_FIELDS macro.

void TARGET_MACHINE_DEPENDENT_REORG (void) [Target Hook]
If non-null, this hook performs a target-specific pass over the instruction stream. The compiler will run it at all optimization levels, just before the point at which it normally does delayed-branch scheduling.

The exact purpose of the hook varies from target to target. Some use it to do transformations that are necessary for correctness, such as laying out in-function constant pools or avoiding hardware hazards. Others use it as an opportunity to do some machine-dependent optimizations.

You need not implement the hook if it has nothing to do. The default definition is null.

void TARGET_INIT_BUILTINS (void) [Target Hook]
Define this hook if you have any machine-specific built-in functions that need to be defined. It should be a function that performs the necessary setup.

Machine specific built-in functions can be useful to expand special machine instructions that would otherwise not normally be generated because they have no equivalent in the source language (for example, SIMD vector instructions or prefetch instructions).

To create a built-in function, call the function lang_hooks.builtin_function which is defined by the language front end. You can use any type nodes set up by build_common_tree_nodes and build_common_tree_nodes_2; only language front ends that use those two functions will call 'TARGET_INIT_BUILTINS'.
tree TARGET_BUILTIN_DECL (unsigned code, bool initialize_p)  [Target Hook]
Define this hook if you have any machine-specific built-in functions that need to be defined. It should be a function that returns the built-in function declaration for the built-in function code code. If there is no such builtin and it cannot be initialized at this time if initialize_p is true the function should return NULL_TREE. If code is out of range the function should return error_mark_node.

rtx TARGET_EXPAND_BUILTIN (tree exp, rtx target, rtx
   subtarget, enum machine_mode mode, int ignore)  [Target Hook]
Expand a call to a machine specific built-in function that was set up by
   ‘TARGET_INIT_BUILTIN’. exp is the expression for the function call; the result
should go to target if that is convenient, and have mode mode if that is convenient.
   subtarget may be used as the target for computing one of exp’s operands. ignore is nonzero if the value is to be ignored. This function should return the result of the call to the built-in function.

tree TARGET_RESOLVE_OVERLOADED_BUILTIN (unsigned int loc,
   tree fndecl, void *arglist)  [Target Hook]
Select a replacement for a machine specific built-in function that was set up by
   ‘TARGET_INIT_BUILTIN’. This is done before regular type checking, and so allows
the target to implement a crude form of function overloading. fndecl is the declaration of the built-in function. arglist is the list of arguments passed to the built-in function. The result is a complete expression that implements the operation, usually another CALL_EXPR. arglist really has type ‘VEC(tree,gc)*’

tree TARGET_FOLD_BUILTIN (tree fndecl, int n_args, tree *argp,
   bool ignore)  [Target Hook]
Fold a call to a machine specific built-in function that was set up by
   ‘TARGET_INIT_BUILTIN’. fndecl is the declaration of the built-in function. n_args
is the number of arguments passed to the function; the arguments themselves are
pointed to by argp. The result is another tree containing a simplified expression for
the call’s result. If ignore is true the value will be ignored.

const char * TARGET_INVALID_WITHIN_DOLOOP (const rtx insn)  [Target Hook]
Take an instruction in insn and return NULL if it is valid within a low-overhead loop,
otherwise return a string explaining why doloop could not be applied.
Many targets use special registers for low-overhead looping. For any instruction that
cllobers these this function should return a string indicating the reason why the
doloop could not be applied. By default, the RTL loop optimizer does not use a
present doloop pattern for loops containing function calls or branch on table instruc-
tions.

MD_CAN_REDIRECT_BRANCH (branch1, branch2)  [Macro]
Take a branch insn in branch1 and another in branch2. Return true if redirecting
branch1 to the destination of branch2 is possible.
On some targets, branches may have a limited range. Optimizing the filling of delay
slots can result in branches being redirected, and this may in turn cause a branch
offset to overflow.
bool TARGET_COMMUTATIVE_P (const rtx x, int outer_code)  [Target Hook]
This target hook returns true if x is considered to be commutative. Usually, this
is just COMMUTATIVE_P (x), but the HP PA doesn’t consider PLUS to be com-
mutable inside a MEM. outer_code is the rtx code of the enclosing rtl, if known,
otherwise it is UNKNOWN.

rtx TARGET_ALLOCATE_INITIAL_VALUE (rtx hard_reg)  [Target Hook]
When the initial value of a hard register has been copied in a pseudo register, it
is often not necessary to actually allocate another register to this pseudo register,
because the original hard register or a stack slot has been saved into can be used.
TARGET_ALLOCATE_INITIAL_VALUE is called at the start of register allocation once
for each hard register that had its initial value copied by using get_func_hard_reg_
initial_val or get_hard_reg_initial_val. Possible values are NULL_RTX, if you
don’t want to do any special allocation, a REG rtx—that would typically be the hard
register itself, if it is known not to be clobbered—or a MEM. If you are returning a MEM,
this is only a hint for the allocator; it might decide to use another register anyways.
You may use current_function_leaf_function in the hook, functions that use
REG_N_SETS, to determine if the hard register in question will not be clobbered. The
default value of this hook is NULL, which disables any special allocation.

int TARGET_UNSPEC_MAY_TRAP_P (const rtx x, unsigned flags)  [Target Hook]
This target hook returns nonzero if x, an unspec or unspec_volatile operation,
might cause a trap. Targets can use this hook to enhance precision of analysis for
unspec and unspec_volatile operations. You may call may_trap_p_1 to analyze
inner elements of x in which case flags should be passed along.

void TARGET_SET_CURRENT_FUNCTION (tree decl)  [Target Hook]
The compiler invokes this hook whenever it changes its current function context
(cfun). You can define this function if the back end needs to perform any initial-
ization or reset actions on a per-function basis. For example, it may be used to
implement function attributes that affect register usage or code generation patterns.
The argument decl is the declaration for the new function context, and may be null
to indicate that the compiler has left a function context and is returning to processing
at the top level. The default hook function does nothing.
GCC sets cfun to a dummy function context during initialization of some parts of
the back end. The hook function is not invoked in this situation; you need not worry
about the hook being invoked recursively, or when the back end is in a partially-
initialized state. cfun might be NULL to indicate processing at top level, outside of
any function scope.

TARGET_OBJECT_SUFFIX  [Macro]
Define this macro to be a C string representing the suffix for object files on your
target machine. If you do not define this macro, GCC will use ‘.o’ as the suffix for
object files.

TARGET_EXECUTABLE_SUFFIX  [Macro]
Define this macro to be a C string representing the suffix to be automatically added
to executable files on your target machine. If you do not define this macro, GCC will
use the null string as the suffix for executable files.
COLLECT_EXPORT_LIST [Macro]
If defined, collect2 will scan the individual object files specified on its command line and create an export list for the linker. Define this macro for systems like AIX, where the linker discards object files that are not referenced from main and uses export lists.

MODIFY_JNI_METHOD_CALL (mdecl) [Macro]
Define this macro to a C expression representing a variant of the method call mdecl, if Java Native Interface (JNI) methods must be invoked differently from other methods on your target. For example, on 32-bit Microsoft Windows, JNI methods must be invoked using the stdcall calling convention and this macro is then defined as this expression:

\[
\text{build_type_attribute_variant (mdecl, build_tree_list (get_identifier ("stdcall"), NULL))}
\]

bool TARGET_CANNOT_MODIFY_JUMPS_P (void) [Target Hook]
This target hook returns true past the point in which new jump instructions could be created. On machines that require a register for every jump such as the SHmedia ISA of SH5, this point would typically be reload, so this target hook should be defined to a function such as:

```c
static bool cannot_modify_jumps_past_reload_p ()
{
    return (reload_completed || reload_in_progress);
}
```

reg_class_t TARGET_BRANCH_TARGET_REGISTER_CLASS (void) [Target Hook]
This target hook returns a register class for which branch target register optimizations should be applied. All registers in this class should be usable interchangeably. After reload, registers in this class will be re-allocated and loads will be hoisted out of loops and be subjected to inter-block scheduling.

bool TARGET_BRANCH_TARGET_REGISTER_CALLEE_SAVED (bool after_prologue_epilogue_gen) [Target Hook]
Branch target register optimization will by default exclude callee-saved registers that are not already live during the current function; if this target hook returns true, they will be included. The target code must than make sure that all target registers in the class returned by ‘TARGET_BRANCH_TARGET_REGISTER_CLASS’ that might need saving are saved. after_prologue_epilogue_gen indicates if prologues and epilogues have already been generated. Note, even if you only return true when after_prologue_epilogue_gen is false, you still are likely to have to make special provisions in INITIAL_ELIMINATION_OFFSET to reserve space for caller-saved target registers.

bool TARGET_HAVE_CONDITIONAL_EXECUTION (void) [Target Hook]
This target hook returns true if the target supports conditional execution. This target hook is required only when the target has several different modes and they have different conditional execution capability, such as ARM.
unsigned TARGET_LOOP_UNROLL_ADJUST (unsigned nunroll, struct loop *loop)  
This target hook returns a new value for the number of times loop should be unrolled. The parameter nunroll is the number of times the loop is to be unrolled. The parameter loop is a pointer to the loop, which is going to be checked for unrolling. This target hook is required only when the target has special constraints like maximum number of memory accesses.

POWI_MAX_MULTS [Macro]  
If defined, this macro is interpreted as a signed integer C expression that specifies the maximum number of floating point multiplications that should be emitted when expanding exponentiation by an integer constant inline. When this value is defined, exponentiation requiring more than this number of multiplications is implemented by calling the system library’s pow, powf or powl routines. The default value places no upper bound on the multiplication count.

void TARGET_EXTRA_INCLUDES (const char *sysroot, const char *iprefix, int stdinc) [Macro]  
This target hook should register any extra include files for the target. The parameter stdinc indicates if normal include files are present. The parameter sysroot is the system root directory. The parameter iprefix is the prefix for the gcc directory.

void TARGET_EXTRA_PRE_INCLUDES (const char *sysroot, const char *iprefix, int stdinc) [Macro]  
This target hook should register any extra include files for the target before any standard headers. The parameter stdinc indicates if normal include files are present. The parameter sysroot is the system root directory. The parameter iprefix is the prefix for the gcc directory.

void TARGET_OPTF (char *path) [Macro]  
This target hook should register special include paths for the target. The parameter path is the include to register. On Darwin systems, this is used for Framework includes, which have semantics that are different from ‘-I’.

bool TARGET_USE_LOCAL_THUNK_ALIAS_P (tree fndecl) [Macro]  
This target macro returns true if it is safe to use a local alias for a virtual function fndecl when constructing thunks, false otherwise. By default, the macro returns true for all functions, if a target supports aliases (i.e. defines ASM_OUTPUT_DEF), false otherwise,

TARGET_FORMAT_TYPES [Macro]  
If defined, this macro is the name of a global variable containing target-specific format checking information for the ‘-Wformat’ option. The default is to have no target-specific format checks.

TARGET_N_FORMAT_TYPES [Macro]  
If defined, this macro is the number of entries in TARGET_FORMAT_TYPES.
TARGET_OVERRIDES_FORMAT_ATTRIBUTES

If defined, this macro is the name of a global variable containing target-specific format
overrides for the ‘-Wformat’ option. The default is to have no target-specific format
overrides. If defined, TARGET_FORMAT_TYPES must be defined, too.

TARGET_OVERRIDES_FORMAT_ATTRIBUTES_COUNT

If defined, this macro specifies the number of entries in TARGET_OVERRIDES_FORMAT-
ATTRIBUTES.

TARGET_OVERRIDES_FORMAT_INIT

If defined, this macro specifies the optional initialization routine for target specific
customizations of the system printf and scanf formatter settings.

bool TARGET_RELAXED_ORDERING

If set to true, means that the target’s memory model does not guarantee that loads
which do not depend on one another will access main memory in the order of the
instruction stream; if ordering is important, an explicit memory barrier must be
used. This is true of many recent processors which implement a policy of “relaxed,”
“weak,” or “release” memory consistency, such as Alpha, PowerPC, and ia64. The
default is false.

const char * TARGET_INVALID_ARG_FOR_UNPROTOTYPED_FN

(const_tree typelist, const_tree funcdecl, const_tree val)

If defined, this macro returns the diagnostic message when it is illegal to pass argument
val to function funcdecl with prototype typelist.

const char * TARGET_INVALID_CONVERSION

(const_tree fromtype, const_tree totype)

If defined, this macro returns the diagnostic message when it is invalid to convert
from fromtype to totype, or NULL if validity should be determined by the front end.

const char * TARGET_INVALID_UNARY_OP

(int op, const_tree type)

If defined, this macro returns the diagnostic message when it is invalid to apply
operation op (where unary plus is denoted by CONVERT_EXPR) to an operand of type
type, or NULL if validity should be determined by the front end.

const char * TARGET_INVALID_BINARY_OP

(int op, const_tree type1, const_tree type2)

If defined, this macro returns the diagnostic message when it is invalid to apply
operation op to operands of types type1 and type2, or NULL if validity should be
determined by the front end.

const char * TARGET_INVALID_PARAMETER_TYPE

(const_tree type)

If defined, this macro returns the diagnostic message when it is invalid for functions
to include parameters of type type, or NULL if validity should be determined by the
front end. This is currently used only by the C and C++ front ends.
const char * TARGET_INVALID_RETURN_TYPE (const tree type)  [Target Hook]
If defined, this macro returns the diagnostic message when it is invalid for functions
to have return type type, or NULL if validity should be determined by the front end.
This is currently used only by the C and C++ front ends.

tree TARGET_PROMOTED_TYPE (const tree type)  [Target Hook]
If defined, this target hook returns the type to which values of type should be pro-
moted when they appear in expressions, analogous to the integer promotions, or
NULL_TREE to use the front end’s normal promotion rules. This hook is useful when
there are target-specific types with special promotion rules. This is currently used
only by the C and C++ front ends.

tree TARGET_CONVERT_TO_TYPE (tree type, tree expr)  [Target Hook]
If defined, this hook returns the result of converting expr to type. It should return
the converted expression, or NULL_TREE to apply the front end’s normal conversion
rules. This hook is useful when there are target-specific types with special conversion
rules. This is currently used only by the C and C++ front ends.

TARGET_USE_JCR_SECTION  [Macro]
This macro determines whether to use the JCR section to register Java classes. By
default, TARGET_USE_JCR_SECTION is defined to 1 if both SUPPORTS_WEAK
and TARGET_HAVE_NAMED_SECTIONS are true, else 0.

OBJC_JBLEN  [Macro]
This macro determines the size of the objective C jump buffer for the NeXT runtime.
By default, OBJC_JBLEN is defined to an innocuous value.

LIBGCC2_UNWIND_ATTRIBUTE  [Macro]
Define this macro if any target-specific attributes need to be attached to the functions
in ‘libgcc’ that provide low-level support for call stack unwinding. It is used in
declarations in ‘unwind-generic.h’ and the associated definitions of those functions.

void TARGET_UPDATE_STACK_BOUNDARY (void)  [Target Hook]
Define this macro to update the current function stack boundary if necessary.

rtx TARGET_GET_DRAP_RTX (void)  [Target Hook]
This hook should return an rtx for Dynamic Realign Argument Pointer (DRAP) if
a different argument pointer register is needed to access the function’s argument list
due to stack realignment. Return NULL if no DRAP is needed.

bool TARGET_ALLOCATE_STACK_SLOTS_FOR_ARGS (void)  [Target Hook]
When optimization is disabled, this hook indicates whether or not arguments should
be allocated to stack slots. Normally, GCC allocates stacks slots for arguments when
not optimizing in order to make debugging easier. However, when a function is
declared with __attribute__((naked)), there is no stack frame, and the compiler
cannot safely move arguments from the registers in which they are passed to the stack.
Therefore, this hook should return true in general, but false for naked functions. The
default implementation always returns true.
unsigned HOST_WIDE_INT TARGET_CONST_ANCHOR

On some architectures it can take multiple instructions to synthesize a constant. If there is another constant already in a register that is close enough in value then it is preferable that the new constant is computed from this register using immediate addition or subtraction. We accomplish this through CSE. Besides the value of the constant we also add a lower and an upper constant anchor to the available expressions. These are then queried when encountering new constants. The anchors are computed by rounding the constant up and down to a multiple of the value of TARGET_CONST_ANCHOR. TARGET_CONST_ANCHOR should be the maximum positive value accepted by immediate-add plus one. We currently assume that the value of TARGET_CONST_ANCHOR is a power of 2. For example, on MIPS, where add-immediate takes a 16-bit signed value, TARGET_CONST_ANCHOR is set to ‘0x8000’. The default value is zero, which disables this optimization.
18 Host Configuration

Most details about the machine and system on which the compiler is actually running are detected by the configure script. Some things are impossible for configure to detect; these are described in two ways, either by macros defined in a file named ‘xm-machine.h’ or by hook functions in the file specified by the out_host_hook_obj variable in ‘config.gcc’. (The intention is that very few hosts will need a header file but nearly every fully supported host will need to override some hooks.)

If you need to define only a few macros, and they have simple definitions, consider using the xm_defines variable in your ‘config.gcc’ entry instead of creating a host configuration header. See Section 6.3.2.2 [System Config], page 64.

18.1 Host Common

Some things are just not portable, even between similar operating systems, and are too difficult for autoconf to detect. They get implemented using hook functions in the file specified by the host_hook_obj variable in ‘config.gcc’.

void HOST_HOOKS_EXTRA_SIGNALS (void) [Host Hook]
This host hook is used to set up handling for extra signals. The most common thing to do in this hook is to detect stack overflow.

void * HOST_HOOKS_GT_PCH_GET_ADDRESS (size_t size, int fd) [Host Hook]
This host hook returns the address of some space that is likely to be free in some subsequent invocation of the compiler. We intend to load the PCH data at this address such that the data need not be relocated. The area should be able to hold size bytes. If the host uses mmap, fd is an open file descriptor that can be used for probing.

int HOST_HOOKS_GT_PCH_USE_ADDRESS (void * address, size_t size, int fd, size_t offset) [Host Hook]
This host hook is called when a PCH file is about to be loaded. We want to load size bytes from fd at offset into memory at address. The given address will be the result of a previous invocation of HOST_HOOKS_GT_PCH_GET_ADDRESS. Return −1 if we couldn’t allocate size bytes at address. Return 0 if the memory is allocated but the data is not loaded. Return 1 if the hook has performed everything.

If the implementation uses reserved address space, free any reserved space beyond size, regardless of the return value. If no PCH will be loaded, this hook may be called with size zero, in which case all reserved address space should be freed.

Do not try to handle values of address that could not have been returned by this executable; just return −1. Such values usually indicate an out-of-date PCH file (built by some other GCC executable), and such a PCH file won’t work.

size_t HOST_HOOKS_GT_PCH_ALLOC_GRANULARITY (void); [Host Hook]
This host hook returns the alignment required for allocating virtual memory. Usually this is the same as getpagesize, but on some hosts the alignment for reserving memory differs from the pagesize for committing memory.
18.2 Host Filesystem

GCC needs to know a number of things about the semantics of the host machine’s filesystem. Filesystems with Unix and MS-DOS semantics are automatically detected. For other systems, you can define the following macros in ‘xm-machine.h’.

HAVE_DOS_BASED_FILE_SYSTEM
This macro is automatically defined by ‘system.h’ if the host file system obeys the semantics defined by MS-DOS instead of Unix. DOS file systems are case insensitive, file specifications may begin with a drive letter, and both forward slash and backslash (’/’ and ’\’) are directory separators.

DIR_SEPARATOR
DIR_SEPARATOR_2
If defined, these macros expand to character constants specifying separators for directory names within a file specification. ‘system.h’ will automatically give them appropriate values on Unix and MS-DOS file systems. If your file system is neither of these, define one or both appropriately in ‘xm-machine.h’.

However, operating systems like VMS, where constructing a pathname is more complicated than just stringing together directory names separated by a special character, should not define either of these macros.

PATH_SEPARATOR
If defined, this macro should expand to a character constant specifying the separator for elements of search paths. The default value is a colon (‘:’). DOS-based systems usually, but not always, use semicolon (‘;’).

VMS
Define this macro if the host system is VMS.

HOST_OBJECT_SUFFIX
Define this macro to be a C string representing the suffix for object files on your host machine. If you do not define this macro, GCC will use ‘.o’ as the suffix for object files.

HOST_EXECUTABLE_SUFFIX
Define this macro to be a C string representing the suffix for executable files on your host machine. If you do not define this macro, GCC will use the null string as the suffix for executable files.

HOST_BIT_BUCKET
A pathname defined by the host operating system, which can be opened as a file and written to, but all the information written is discarded. This is commonly known as a bit bucket or null device. If you do not define this macro, GCC will use ‘/dev/null’ as the bit bucket. If the host does not support a bit bucket, define this macro to an invalid filename.

UPDATE_PATH_HOST_CANONICALIZE (path)
If defined, a C statement (sans semicolon) that performs host-dependent canonicalization when a path used in a compilation driver or preprocessor is canonicalized. path is a malloc-ed path to be canonicalized. If the C statement does canonicalize path into a different buffer, the old path should be freed and the new buffer should have been allocated with malloc.
DUMPFILE_FORMAT
Define this macro to be a C string representing the format to use for constructing the index part of debugging dump file names. The resultant string must fit in fifteen bytes. The full filename will be the concatenation of: the prefix of the assembler file name, the string resulting from applying this format to an index number, and a string unique to each dump file kind, e.g. ‘rtl’.
If you do not define this macro, GCC will use ‘.%02d.’. You should define this macro if using the default will create an invalid file name.

DELETE_IF_ORDINARY
Define this macro to be a C statement (sans semicolon) that performs host-dependent removal of ordinary temp files in the compilation driver.
If you do not define this macro, GCC will use the default version. You should define this macro if the default version does not reliably remove the temp file as, for example, on VMS which allows multiple versions of a file.

HOST_LACKS_INODE_NUMBERS
Define this macro if the host filesystem does not report meaningful inode numbers in struct stat.

18.3 Host Misc

FATAL_EXIT_CODE
A C expression for the status code to be returned when the compiler exits after serious errors. The default is the system-provided macro ‘EXIT_FAILURE’, or ‘1’ if the system doesn’t define that macro. Define this macro only if these defaults are incorrect.

SUCCESS_EXIT_CODE
A C expression for the status code to be returned when the compiler exits without serious errors. (Warnings are not serious errors.) The default is the system-provided macro ‘EXIT_SUCCESS’, or ‘0’ if the system doesn’t define that macro. Define this macro only if these defaults are incorrect.

USE_C_ALLOCA
Define this macro if GCC should use the C implementation of alloca provided by ‘libiberty.a’. This only affects how some parts of the compiler itself allocate memory. It does not change code generation.
When GCC is built with a compiler other than itself, the C alloca is always used. This is because most other implementations have serious bugs. You should define this macro only on a system where no stack-based alloca can possibly work. For instance, if a system has a small limit on the size of the stack, GCC’s builtin alloca will not work reliably.

COLLECT2_HOST_INITIALIZATION
If defined, a C statement (sans semicolon) that performs host-dependent initialization when collect2 is being initialized.

GCC_DRIVER_HOST_INITIALIZATION
If defined, a C statement (sans semicolon) that performs host-dependent initialization when a compilation driver is being initialized.
HOST_LONG_LONG_FORMAT
If defined, the string used to indicate an argument of type long long to functions like printf. The default value is "ll".

HOST_LONG_FORMAT
If defined, the string used to indicate an argument of type long to functions like printf. The default value is "l".

HOST_PTR_PRINTF
If defined, the string used to indicate an argument of type void * to functions like printf. The default value is "%p".

In addition, if configure generates an incorrect definition of any of the macros in ‘auto-host.h’, you can override that definition in a host configuration header. If you need to do this, first see if it is possible to fix configure.
19 Makefile Fragments

When you configure GCC using the `configure` script, it will construct the file `Makefile` from the template file `Makefile.in`. When it does this, it can incorporate makefile fragments from the `config` directory. These are used to set Makefile parameters that are not amenable to being calculated by autoconf. The list of fragments to incorporate is set by `config.gcc` (and occasionally `config.build` and `config.host`); See Section 6.3.2.2 [System Config], page 64.

Fragments are named either `t-target` or `x-host`, depending on whether they are relevant to configuring GCC to produce code for a particular target, or to configuring GCC to run on a particular host. Here `target` and `host` are mnemonics which usually have some relationship to the canonical system name, but no formal connection.

If these files do not exist, it means nothing needs to be added for a given target or host. Most targets need a few `t-target` fragments, but needing `x-host` fragments is rare.

19.1 Target Makefile Fragments

Target makefile fragments can set these Makefile variables.

`LIBGCC2_CFLAGS`

Compiler flags to use when compiling `libgcc2.c`.

`LIB2FUNCS_EXTRA`

A list of source file names to be compiled or assembled and inserted into `libgcc.a`.

Floating Point Emulation

To have GCC include software floating point libraries in `libgcc.a` define `FPBIT` and `DPBIT` along with a few rules as follows:

```
# We want fine grained libraries, so use the new code
# to build the floating point emulation libraries.
FPBIT = fp-bit.c
DPBIT = dp-bit.c
```

```
fp-bit.c: $(srcdir)/config/fp-bit.c
    echo '#define FLOAT' > fp-bit.c
    cat $(srcdir)/config/fp-bit.c >> fp-bit.c

dp-bit.c: $(srcdir)/config/fp-bit.c
    cat $(srcdir)/config/fp-bit.c > dp-bit.c
```

You may need to provide additional #defines at the beginning of `fp-bit.c` and `dp-bit.c` to control target endianness and other options.

`CRTSTUFF_T_CFLAGS`

Special flags used when compiling `crtstuff.c`. See Section 17.21.5 [Initialization], page 512.

`CRTSTUFF_T_CFLAGS_S`

Special flags used when compiling `crtstuff.c` for shared linking. Used if you use `crtbeginS.o` and `crtendS.o` in `EXTRA-PARTS`. See Section 17.21.5 [Initialization], page 512.
MULTILIB_OPTIONS

For some targets, invoking GCC in different ways produces objects that cannot be linked together. For example, for some targets GCC produces both big and little endian code. For these targets, you must arrange for multiple versions of ‘libgcc.a’ to be compiled, one for each set of incompatible options. When GCC invokes the linker, it arranges to link in the right version of ‘libgcc.a’, based on the command line options used.

The MULTILIB_OPTIONS macro lists the set of options for which special versions of ‘libgcc.a’ must be built. Write options that are mutually incompatible side by side, separated by a slash. Write options that may be used together separated by a space. The build procedure will build all combinations of compatible options.

For example, if you set MULTILIB_OPTIONS to ‘m68000/m68020 msoft-float’, ‘Makefile’ will build special versions of ‘libgcc.a’ using the following sets of options: ‘-m68000’, ‘-m68020’, ‘-msoft-float’, ‘-m68000 -msoft-float’, and ‘-m68020 -msoft-float’.

MULTILIB_DIRNAMES

If MULTILIB_OPTIONS is used, this variable specifies the directory names that should be used to hold the various libraries. Write one element in MULTILIB_DIRNAMES for each element in MULTILIB_OPTIONS. If MULTILIB_DIRNAMES is not used, the default value will be MULTILIB_OPTIONS, with all slashes treated as spaces.

For example, if MULTILIB_OPTIONS is set to ‘m68000/m68020 msoft-float’, then the default value of MULTILIB_DIRNAMES is ‘m68000 m68020 msoft-float’. You may specify a different value if you desire a different set of directory names.

MULTILIB_MATCHES

Sometimes the same option may be written in two different ways. If an option is listed in MULTILIB_OPTIONS, GCC needs to know about any synonyms. In that case, set MULTILIB_MATCHES to a list of items of the form ‘option=option’ to describe all relevant synonyms. For example, ‘m68000=mc68000 m68020=mc68020’.

MULTILIB_EXCEPTIONS

Sometimes when there are multiple sets of MULTILIB_OPTIONS being specified, there are combinations that should not be built. In that case, set MULTILIB_EXCEPTIONS to be all of the switch exceptions in shell case syntax that should not be built.

For example the ARM processor cannot execute both hardware floating point instructions and the reduced size THUMB instructions at the same time, so there is no need to build libraries with both of these options enabled. Therefore MULTILIB_EXCEPTIONS is set to:

```
*mthumb/*mhard-float*
```

MULTILIB_EXTRA_OPTS

Sometimes it is desirable that when building multiple versions of ‘libgcc.a’ certain options should always be passed on to the compiler. In that case, set MULTILIB_EXTRA_OPTS to be the list of options to be used for all builds. If you set this, you should probably set CRTSTUFF_T_CFLAGS to a dash followed by it.
NATIVE_SYSTEM_HEADER_DIR
If the default location for system headers is not ‘/usr/include’, you must set this to the directory containing the headers. This value should match the value of the SYSTEM_INCLUDE_DIR macro.

SPECS
Unfortunately, setting MULTILIB_EXTRA_OPTS is not enough, since it does not affect the build of target libraries, at least not the build of the default multilib. One possible work-around is to use DRIVER_SELF_SPECS to bring options from the ‘specs’ file as if they had been passed in the compiler driver command line. However, you don’t want to be adding these options after the toolchain is installed, so you can instead tweak the ‘specs’ file that will be used during the toolchain build, while you still install the original, built-in ‘specs’. The trick is to set SPECS to some other filename (say ‘specs.install’), that will then be created out of the built-in specs, and introduce a ‘Makefile’ rule to generate the ‘specs’ file that’s going to be used at build time out of your ‘specs.install’.

T_CFLAGS
These are extra flags to pass to the C compiler. They are used both when building GCC, and when compiling things with the just-built GCC. This variable is deprecated and should not be used.

19.2 Host Makefile Fragments
The use of ‘x-host’ fragments is discouraged. You should only use it for makefile dependencies.
20 collect2

GCC uses a utility called collect2 on nearly all systems to arrange to call various initialization functions at start time.

The program collect2 works by linking the program once and looking through the linker output file for symbols with particular names indicating they are constructor functions. If it finds any, it creates a new temporary `.c` file containing a table of them, compiles it, and links the program a second time including that file.

The actual calls to the constructors are carried out by a subroutine called `__main`, which is called (automatically) at the beginning of the body of `main` (provided `main` was compiled with GNU CC). Calling `__main` is necessary, even when compiling C code, to allow linking C and C++ object code together. (If you use `--nostdlib`, you get an unresolved reference to `__main`, since it’s defined in the standard GCC library. Include `--lgcc` at the end of your compiler command line to resolve this reference.)

The program collect2 is installed as `ld` in the directory where the passes of the compiler are installed. When collect2 needs to find the `real ld`, it tries the following file names:

- a hard coded linker file name, if GCC was configured with the `--with-ld` option.
- `real-ld` in the directories listed in the compiler’s search directories.
- `real-ld` in the directories listed in the environment variable `PATH`.
- The file specified in the `REAL_LD_FILE_NAME` configuration macro, if specified.
- `ld` in the compiler’s search directories, except that collect2 will not execute itself recursively.
- `ld` in `PATH`.

“The compiler’s search directories” means all the directories where gcc searches for passes of the compiler. This includes directories that you specify with `--B`.

Cross-compilers search a little differently:

- `real-ld` in the compiler’s search directories.
- `target-real-ld` in `PATH`.
- The file specified in the `REAL_LD_FILE_NAME` configuration macro, if specified.
- `ld` in the compiler’s search directories.
- `target-ld` in `PATH`.

collect2 explicitly avoids running ld using the file name under which collect2 itself was invoked. In fact, it remembers up a list of such names—in case one copy of collect2 finds another copy (or version) of collect2 installed as ld in a second place in the search path.

collect2 searches for the utilities `nm` and `strip` using the same algorithm as above for ld.
21 Standard Header File Directories

GCC\_INCLUDE\_DIR means the same thing for native and cross. It is where GCC stores its private include files, and also where GCC stores the fixed include files. A cross compiled GCC runs \texttt{fixincludes} on the header files in ‘$\texttt{(tooldir)/include}$’. (If the cross compilation header files need to be fixed, they must be installed before GCC is built. If the cross compilation header files are already suitable for GCC, nothing special need be done).

GPLUSPLUS\_INCLUDE\_DIR means the same thing for native and cross. It is where \texttt{g++} looks first for header files. The C++ library installs only target independent header files in that directory.

LOCAL\_INCLUDE\_DIR is used only by native compilers. GCC doesn’t install anything there. It is normally ‘/usr/local/include’. This is where local additions to a packaged system should place header files.

CROSS\_INCLUDE\_DIR is used only by cross compilers. GCC doesn’t install anything there.

TOOL\_INCLUDE\_DIR is used for both native and cross compilers. It is the place for other packages to install header files that GCC will use. For a cross-compiler, this is the equivalent of ‘/usr/include’. When you build a cross-compiler, \texttt{fixincludes} processes any header files in this directory.
22 Memory Management and Type Information

GCC uses some fairly sophisticated memory management techniques, which involve determining information about GCC’s data structures from GCC’s source code and using this information to perform garbage collection and implement precompiled headers.

A full C parser would be too complicated for this task, so a limited subset of C is interpreted and special markers are used to determine what parts of the source to look at. All struct and union declarations that define data structures that are allocated under control of the garbage collector must be marked. All global variables that hold pointers to garbage-collected memory must also be marked. Finally, all global variables that need to be saved and restored by a precompiled header must be marked. (The precompiled header mechanism can only save static variables if they’re scalar. Complex data structures must be allocated in garbage-collected memory to be saved in a precompiled header.)

The full format of a marker is

```
GTY ((option) [(param)], [option] [(param)] ...)
```

but in most cases no options are needed. The outer double parentheses are still necessary, though: GTY(()). Markers can appear:

- In a structure definition, before the open brace;
- In a global variable declaration, after the keyword static or extern; and
- In a structure field definition, before the name of the field.

Here are some examples of marking simple data structures and globals.

```c
struct GTY(()) tag
{
    fields...
};

typedef struct GTY(()) tag
{
    fields...
} *typename;

static GTY(()) struct tag *list; /* points to GC memory */
static GTY(()) int counter; /* save counter in a PCH */
```

The parser understands simple typedefs such as `typedef struct tag *name;` and `typedef int name;`. These don’t need to be marked.

22.1 The Inside of a GTY(())

Sometimes the C code is not enough to fully describe the type structure. Extra information can be provided with GTY options and additional markers. Some options take a parameter, which may be either a string or a type name, depending on the parameter. If an option takes no parameter, it is acceptable either to omit the parameter entirely, or to provide an empty string as a parameter. For example, `GTY ((skip))` and `GTY ((skip ("")))` are equivalent.

When the parameter is a string, often it is a fragment of C code. Four special escapes of C code, to refer to pieces of the data structure being marked:

```c
%h  The current structure.
```
The structure that immediately contains the current structure.

The outermost structure that contains the current structure.

A partial expression of the form \([i_1][i_2]\ldots\) that indexes the array item currently being marked.

For instance, suppose that you have a structure of the form
\[
\begin{aligned}
\text{struct } A & \{ \\
& \ldots \\
\}\; \\
\text{struct } B & \{ \\
& \text{struct } A \text{ foo}[12]; \\
\}\;
\end{aligned}
\]
and \(b\) is a variable of type \(\text{struct } B\). When marking \(\text{`b.foo[11]'}\), %h would expand to \(\text{`b.foo[11]'}\), %0 and %1 would both expand to \(\text{`b'}\), and %a would expand to \(\text{`[11]'}\).

As in ordinary C, adjacent strings will be concatenated; this is helpful when you have a complicated expression.

\[
\text{GTY} \left(\text{chain\_next \left(\text{"TREE\_CODE } (&%h.\text{generic}) == INTEGER\_TYPE"} \right.\right.
\begin{aligned}
& \left.\text{" ? TYPE\_NEXT\_VARIANT } (\&%h.\text{generic})\text{"}
& \left.\text{" : TREE\_CHAIN } (\&%h.\text{generic})\text{"})\right)\right)
\]

The available options are:

**length ("expression")**

There are two places the type machinery will need to be explicitly told the length of an array. The first case is when a structure ends in a variable-length array, like this:

\[
\begin{aligned}
\text{struct } \text{GTY(()) } \text{rtvec\_def } \{ \\
& \text{int } \text{num\_elem};\quad \text{/* number of elements */} \\
& \text{rtx } \text{GTY( (\text{length } (\"%h.\text{num\_elem}\"))) elem[1];} \\
\}\;
\end{aligned}
\]

In this case, the **length** option is used to override the specified array length (which should usually be 1). The parameter of the option is a fragment of C code that calculates the length.

The second case is when a structure or a global variable contains a pointer to an array, like this:

\[
\begin{aligned}
\text{struct } \text{gimple\_omp\_for\_iter } \ast \text{GTY((\text{length } (\"%h.\text{collapse}\"))) iter;}
\end{aligned}
\]

In this case, **iter** has been allocated by writing something like

\[
\begin{aligned}
x\rightarrow\text{iter } = \text{ggc\_alloc\_cleared\_vec\_gimple\_omp\_for\_iter } (\text{collapse});
\end{aligned}
\]

and the **collapse** provides the length of the field.

This second use of **length** also works on global variables, like:

\[
\begin{aligned}
\text{static } \text{GTY(( (\text{length } (\"\text{reg\_known\_value\_size}\"))) } \text{rtx } \ast \text{reg\_known\_value;}
\end{aligned}
\]

**skip**

If **skip** is applied to a field, the type machinery will ignore it. This is somewhat dangerous; the only safe use is in a union when one field really isn’t ever used.

**desc ("expression")**

**tag ("constant")**

**default**
The type machinery needs to be told which field of a **union** is currently active. This is done by giving each field a constant **tag** value, and then specifying a discriminator using **desc**. The value of the expression given by **desc** is compared against each **tag** value, each of which should be different. If no **tag** is matched, the field marked with **default** is used if there is one, otherwise no field in the union will be marked.

In the **desc** option, the “current structure” is the union that it discriminates. Use `%1` to mean the structure containing it. There are no escapes available to the **tag** option, since it is a constant.

For example,

```c
struct GTY(()) tree_binding {
    struct tree_common common;
    union tree_binding_u {
        tree GTY ((tag ("0"))) scope;
        struct cp_binding_level * GTY ((tag ("1"))) level;
    } GTY ((desc ("BINDING_HAS_LEVEL_P ((tree)&%0)"))) xscope;
    tree value;
};
```

In this example, the value of **BINDING_HAS_LEVEL_P** when applied to a `struct tree_binding *` is presumed to be 0 or 1. If 1, the type mechanism will treat the field **level** as being present and if 0, will treat the field **scope** as being present.

### **param_is** *(type)*

**use_param**

Sometimes it’s convenient to define some data structure to work on generic pointers (that is, **PTR**) and then use it with a specific type. **param_is** specifies the real type pointed to, and **use_param** says where in the generic data structure that type should be put.

For instance, to have a `htab_t` that points to trees, one would write the definition of `htab_t` like this:

```c
typedef struct GTY(()) {
    ...
    void ** GTY ((use_param, ...)) entries;
    ...
} htab_t;
```

and then declare variables like this:

```c
static htab_t GTY ((param_is (union tree_node))) ict;
```

### **paramn_is** *(type)*

**use_paramn**

In more complicated cases, the data structure might need to work on several different types, which might not necessarily all be pointers. For this, **param1_is** through **param9_is** may be used to specify the real type of a field identified by **use_param1** through **use_param9**.

**use_params**

When a structure contains another structure that is parameterized, there’s no need to do anything special, the inner structure inherits the parameters of the
outer one. When a structure contains a pointer to a parameterized structure, the type machinery won’t automatically detect this (it could, it just doesn’t yet), so it’s necessary to tell it that the pointed-to structure should use the same parameters as the outer structure. This is done by marking the pointer with the `use_params` option.

**deletable**

`deletable`, when applied to a global variable, indicates that when garbage collection runs, there’s no need to mark anything pointed to by this variable, it can just be set to `NULL` instead. This is used to keep a list of free structures around for re-use.

**if_marked** ("expression")

Suppose you want some kinds of object to be unique, and so you put them in a hash table. If garbage collection marks the hash table, these objects will never be freed, even if the last other reference to them goes away. GGC has special handling to deal with this: if you use the `if_marked` option on a global hash table, GGC will call the routine whose name is the parameter to the option on each hash table entry. If the routine returns nonzero, the hash table entry will be marked as usual. If the routine returns zero, the hash table entry will be deleted.

The routine `ggc_marked_p` can be used to determine if an element has been marked already; in fact, the usual case is to use `if_marked ("ggc_marked_p")`.

**mark_hook** ("hook-routine-name")

If provided for a structure or union type, the given `hook-routine-name` (between double-quotes) is the name of a routine called when the garbage collector has just marked the data as reachable. This routine should not change the data, or call any ggc routine. Its only argument is a pointer to the just marked (const) structure or union.

**maybe_undef**

When applied to a field, `maybe_undef` indicates that it’s OK if the structure that this fields points to is never defined, so long as this field is always `NULL`. This is used to avoid requiring backends to define certain optional structures. It doesn’t work with language frontends.

**nested_ptr** (type, "to expression", "from expression")

The type machinery expects all pointers to point to the start of an object. Sometimes for abstraction purposes it’s convenient to have a pointer which points inside an object. So long as it’s possible to convert the original object to and from the pointer, such pointers can still be used. `type` is the type of the original object, the `to expression` returns the pointer given the original object, and the `from expression` returns the original object given the pointer. The pointer will be available using the `%h` escape.

**chain_next** ("expression")

**chain_prev** ("expression")

**chain_circular** ("expression")

It’s helpful for the type machinery to know if objects are often chained together in long lists; this lets it generate code that uses less stack space by iterating
along the list instead of recursing down it. \texttt{chain\_next} is an expression for the next item in the list, \texttt{chain\_prev} is an expression for the previous item. For singly linked lists, use only \texttt{chain\_next}; for doubly linked lists, use both. The machinery requires that taking the next item of the previous item gives the original item. \texttt{chain\_circular} is similar to \texttt{chain\_next}, but can be used for circular single linked lists.

\texttt{reorder ("function name")}

Some data structures depend on the relative ordering of pointers. If the precompiled header machinery needs to change that ordering, it will call the function referenced by the \texttt{reorder} option, before changing the pointers in the object that's pointed to by the field the option applies to. The function must take four arguments, with the signature \texttt{‘void *, void *, gt\_pointer\_operator, void *’}. The first parameter is a pointer to the structure that contains the object being updated, or the object itself if there is no containing structure. The second parameter is a cookie that should be ignored. The third parameter is a routine that, given a pointer, will update it to its correct new value. The fourth parameter is a cookie that must be passed to the second parameter.

PCH cannot handle data structures that depend on the absolute values of pointers. \texttt{reorder} functions can be expensive. When possible, it is better to depend on properties of the data, like an ID number or the hash of a string instead.

\texttt{variable\_size}

The type machinery expects the types to be of constant size. When this is not true, for example, with structs that have array fields or unions, the type machinery cannot tell how many bytes need to be allocated at each allocation. The \texttt{variable\_size} is used to mark such types. The type machinery then provides allocators that take a parameter indicating an exact size of object being allocated. Note that the size must be provided in bytes whereas the \texttt{length} option works with array lengths in number of elements.

For example,

\begin{verbatim}
struct GTY((variable_size)) sorted_fields_type {
    int len;
    tree GTY((length (“%h.len”))) elts[1];
};
\end{verbatim}

Then the objects of \texttt{struct sorted\_fields\_type} are allocated in GC memory as follows:

\begin{verbatim}
field_vec = ggc_alloc_sorted_fields_type (size);
\end{verbatim}

If \texttt{field\_vec->elts} stores \texttt{n} elements, then \texttt{size} could be calculated as follows:

\begin{verbatim}
size_t size = sizeof (struct sorted\_fields\_type) + n * sizeof (tree);
\end{verbatim}

\texttt{special ("name")}

The \texttt{special} option is used to mark types that have to be dealt with by special case machinery. The parameter is the name of the special case. See ‘\texttt{gengtype.c}’ for further details. Avoid adding new special cases unless there is no other alternative.
22.2 Marking Roots for the Garbage Collector

In addition to keeping track of types, the type machinery also locates the global variables (roots) that the garbage collector starts at. Roots must be declared using one of the following syntaxes:

- `extern GTY( [[options]] ) type name;`
- `static GTY( [[options]] ) type name;`

The syntax

- `GTY( [[options]] ) type name;`

is not accepted. There should be an `extern` declaration of such a variable in a header somewhere—mark that, not the definition. Or, if the variable is only used in one file, make it `static`.

22.3 Source Files Containing Type Information

Whenever you add GTY markers to a source file that previously had none, or create a new source file containing GTY markers, there are three things you need to do:

1. You need to add the file to the list of source files the type machinery scans. There are four cases:
   a. For a back-end file, this is usually done automatically; if not, you should add it to `target_gtfiles` in the appropriate port’s entries in `config.gcc`.
   b. For files shared by all front ends, add the filename to the `GTFILES` variable in `Makefile.in`.
   c. For files that are part of one front end, add the filename to the `gtfiles` variable defined in the appropriate `config-lang.in`. For C, the file is `c-config-lang.in`. Headers should appear before non-headers in this list.
   d. For files that are part of some but not all front ends, add the filename to the `gtfiles` variable of all the front ends that use it.

2. If the file was a header file, you’ll need to check that it’s included in the right place to be visible to the generated files. For a back-end header file, this should be done automatically. For a front-end header file, it needs to be included by the same file that includes `gtype-lang.h`. For other header files, it needs to be included in ‘gtype-desc.c’, which is a generated file, so add it to `ifiles` in `open_base_file` in `gengtype.c`.

For source files that aren’t header files, the machinery will generate a header file that should be included in the source file you just changed. The file will be called `gt-path.h` where `path` is the pathname relative to the gcc directory with slashes replaced by -, so for example the header file to be included in ‘cp/parser.c’ is called `gt-cp-parser.c`. The generated header file should be included after everything else in the source file. Don’t forget to mention this file as a dependency in the `Makefile`!

For language frontends, there is another file that needs to be included somewhere. It will be called `gtype-lang.h`, where `lang` is the name of the subdirectory the language is contained in.

Plugins can add additional root tables. Run the `gengtype` utility in plugin mode as `gengtype -P pluginout.h source-dir file-list plugin*.c` with your plugin files plu-
gin*.c using GTY to generate the pluginout.h file. The GCC build tree is needed to be present in that mode.

22.4 How to invoke the garbage collector

The GCC garbage collector GGC is only invoked explicitly. In contrast with many other garbage collectors, it is not implicitly invoked by allocation routines when a lot of memory has been consumed. So the only way to have GGC reclaim storage it to call the ggc_collect function explicitly. This call is an expensive operation, as it may have to scan the entire heap. Beware that local variables (on the GCC call stack) are not followed by such an invocation (as many other garbage collectors do): you should reference all your data from static or external GTY-ed variables, and it is advised to call ggc_collect with a shallow call stack. The GGC is an exact mark and sweep garbage collector (so it does not scan the call stack for pointers). In practice GCC passes don’t often call ggc_collect themselves, because it is called by the pass manager between passes.

At the time of the ggc_collect call all pointers in the GC-marked structures must be valid or NULL. In practice this means that there should not be uninitialized pointer fields in the structures even if your code never reads or writes those fields at a particular instance. One way to ensure this is to use cleared versions of allocators unless all the fields are initialized manually immediately after allocation.

22.5 Troubleshooting the garbage collector

With the current garbage collector implementation, most issues should show up as GCC compilation errors. Some of the most commonly encountered issues are described below.

- Gengtype does not produce allocators for a GTY-marked type. Gengtype checks if there is at least one possible path from GC roots to at least one instance of each type before outputting allocators. If there is no such path, the GTY markers will be ignored and no allocators will be output. Solve this by making sure that there exists at least one such path. If creating it is unfeasible or raises a “code smell”, consider if you really must use GC for allocating such type.
- Link-time errors about undefined gt_ggc_r_foo_bar and similarly-named symbols. Check if your ‘foo_bar’ source file has #include "gt-foo_bar.h" as its very last line.
Chapter 23: Plugins

23 Plugins

23.1 Loading Plugins

Plugins are supported on platforms that support `-ldl -rdynamic`. They are loaded by the compiler using `dlopen` and invoked at pre-determined locations in the compilation process.

Plugins are loaded with

```
-fplugin=/path/to/name.so -fplugin-arg-name=key1=value1
```

The plugin arguments are parsed by GCC and passed to respective plugins as key-value pairs. Multiple plugins can be invoked by specifying multiple `-fplugin` arguments.

A plugin can be simply given by its short name (no dots or slashes). When simply passing `-fplugin=name`, the plugin is loaded from the `plugin` directory, so `-fplugin=name` is the same as `-fplugin='gcc -print-file-name=plugin'/name.so`, using backquote shell syntax to query the `plugin` directory.

23.2 Plugin API

Plugins are activated by the compiler at specific events as defined in `gcc-plugin.h`. For each event of interest, the plugin should call `register_callback` specifying the name of the event and address of the callback function that will handle that event.

The header `gcc-plugin.h` must be the first gcc header to be included.

23.2.1 Plugin license check

Every plugin should define the global symbol `plugin_is_GPL_compatible` to assert that it has been licensed under a GPL-compatible license. If this symbol does not exist, the compiler will emit a fatal error and exit with the error message:

```
fatal error: plugin name is not licensed under a GPL-compatible license
name: undefined symbol: plugin_is_GPL_compatible
compilation terminated
```

The declared type of the symbol should be int, to match a forward declaration in `gcc-plugin.h` that suppresses C++ mangling. It does not need to be in any allocated section, though. The compiler merely asserts that the symbol exists in the global scope. Something like this is enough:

```
int plugin_is_GPL_compatible;
```

23.2.2 Plugin initialization

Every plugin should export a function called `plugin_init` that is called right after the plugin is loaded. This function is responsible for registering all the callbacks required by the plugin and do any other required initialization.

This function is called from `compile_file` right before invoking the parser. The arguments to `plugin_init` are:

- `plugin_info`: Plugin invocation information.
- `version`: GCC version.

The `plugin_info` struct is defined as follows:
struct plugin_name_args
{
  char *base_name;  /* Short name of the plugin (filename without .so suffix). */
  const char *full_name; /* Path to the plugin as specified with -fplugin=. */
  int argc; /* Number of arguments specified with -fplugin-arg-.... */
  struct plugin_argument *argv; /* Array of ARGC key-value pairs. */
  const char *version; /* Version string provided by plugin. */
  const char *help; /* Help string provided by plugin. */
}

If initialization fails, plugin_init must return a non-zero value. Otherwise, it should return 0.

The version of the GCC compiler loading the plugin is described by the following structure:

struct plugin_gcc_version
{
  const char *basever;
  const char *datestamp;
  const char *devphase;
  const char *revision;
  const char *configuration_arguments;
};

The function plugin_default_version_check takes two pointers to such structure and compare them field by field. It can be used by the plugin’s plugin_init function.

The version of GCC used to compile the plugin can be found in the symbol gcc_version defined in the header 'plugin-version.h'. The recommended version check to perform looks like

#include "plugin-version.h"
...

int plugin_init (struct plugin_name_args *plugin_info,
                 struct plugin_gcc_version *version)
{
  if (!plugin_default_version_check (version, &gcc_version))
    return 1;
}

but you can also check the individual fields if you want a less strict check.

### 23.2.3 Plugin callbacks

Callback functions have the following prototype:

```c
typedef void (*plugin_callback_func)(void *gcc_data, void *user_data);
```

Callbacks can be invoked at the following pre-determined events:

```c
enum plugin_event
{
  PLUGIN_PASS_MANAGER_SETUP,   /* To hook into pass manager. */
```
In addition, plugins can also look up the enumerator of a named event, and / or generate new events dynamically, by calling the function `get_named_event_id`.

To register a callback, the plugin calls `register_callback` with the arguments:

- **char *name**: Plugin name.
- **int event**: The event code.
- **plugin_callback_func callback**: The function that handles event.
- **void *user_data**: Pointer to plugin-specific data.

For the `PLUGIN_PASS_MANAGER_SETUP`, `PLUGIN_INFO`, `PLUGIN_REGISTER_GGC_ROOTS` and `PLUGIN_REGISTER_GGC_CACHES` pseudo-events the callback should be null, and the `user_data` is specific.

When the `PLUGIN_PRAGMAS` event is triggered (with a null pointer as data from GCC), plugins may register their own pragmas using functions like `c_register_pragma` or `c_register_pragma_with_expansion`. 
23.3 Interacting with the pass manager

There needs to be a way to add/reorder/remove passes dynamically. This is useful for both analysis plugins (plugging in after a certain pass such as CFG or an IPA pass) and optimization plugins.

Basic support for inserting new passes or replacing existing passes is provided. A plugin registers a new pass with GCC by calling register_callback with the PLUGIN_PASS_MANAGER_SETUP event and a pointer to a struct register_pass_info object defined as follows:

```c
enum pass_positioning_ops
{
    PASS_POS_INSERT_AFTER, // Insert after the reference pass.
    PASS_POS_INSERT_BEFORE, // Insert before the reference pass.
    PASS_POS_REPLACE       // Replace the reference pass.
};

struct register_pass_info
{
    struct opt_pass *pass;    /* New pass provided by the plugin. */
    const char *reference_pass_name; /* Name of the reference pass for hooking
                                                    up the new pass. */
    int ref_pass_instance_number; /* Insert the pass at the specified
                                               instance number of the reference pass. */
    /* Do it for every instance if it is 0. */
    enum pass_positioning_ops pos_op; /* how to insert the new pass. */
};
```

/* Sample plugin code that registers a new pass. */
int
plugin_init (struct plugin_name_args *plugin_info,
            struct plugin_gcc_version *version)
{
    struct register_pass_info pass_info;

    ...  /* Code to fill in the pass_info object with new pass information. */

    ...  /* Register the new pass. */
    register_callback (plugin_info->base_name, PLUGIN_PASS_MANAGER_SETUP, NULL, &pass_info);

    ...}
```

23.4 Interacting with the GCC Garbage Collector

Some plugins may want to be informed when GGC (the GCC Garbage Collector) is running. They can register callbacks for the PLUGIN_GGC_START and PLUGIN_GGC_END events (for which the callback is called with a null gcc_data) to be notified of the start or end of the GCC garbage collection.

Some plugins may need to have GGC mark additional data. This can be done by registering a callback (called with a null gcc_data) for the PLUGIN_GGC_MARKING event. Such
callbacks can call the `ggc_set_mark` routine, preferably thru the `ggc_mark` macro (and
conversely, these routines should usually not be used in plugins outside of the `PLUGIN_GGC_MARKING` event).

Some plugins may need to add extra GGC root tables, e.g. to handle their own GTY-ed
data. This can be done with the `PLUGIN_REGISTER_GGC_ROOTS` pseudo-event with a null
callback and the extra root table (of type `struct ggc_root_tab*`) as `user_data`. Plugins
that want to use the `if_marked` hash table option can add the extra GGC cache tables
generated by `gengtype` using the `PLUGIN_REGISTER_GGC_CACHES` pseudo-event with a null
callback and the extra cache table (of type `struct ggc_cache_tab*`) as `user_data`. Running the `gengtype -p source-dir file-list plugin*.c ...` utility generates these extra
root tables.

You should understand the details of memory management inside GCC before using
`PLUGIN_GGC_MARKING`, `PLUGIN_REGISTER_GGC_ROOTS` or `PLUGIN_REGISTER_GGC_CACHES`.

### 23.5 Giving information about a plugin

A plugin should give some information to the user about itself. This uses the following
structure:

```c
struct plugin_info
{
    const char *version;
    const char *help;
};
```

Such a structure is passed as the `user_data` by the plugin’s init routine using `register_callback` with the `PLUGIN_INFO` pseudo-event and a null callback.

### 23.6 Registering custom attributes or pragmas

For analysis (or other) purposes it is useful to be able to add custom attributes or pragmas.

The `PLUGIN_ATTRIBUTES` callback is called during attribute registration. Use the
`register_attribute` function to register custom attributes.

```c
/* Attribute handler callback */
static tree
handle_user_attribute (tree *node, tree name, tree args,
    int flags, bool *no_add_attrs)
{
    return NULL_TREE;
}

/* Attribute definition */
static struct attribute_spec user_attr =
    { "user", 1, 1, false, false, false, handle_user_attribute };

/* Plugin callback called during attribute registration.
Registered with register_callback (plugin_name, PLUGIN_ATTRIBUTES, register_attributes, NULL) */
static void
register_attributes (void *event_data, void *data)
{
    warning (0, G_("Callback to register attributes"));
    register_attribute (&user_attr);
}
```
The PLUGIN_PRAGMAS callback is called during pragmas registration. Use the c_register_pragma or c_register_pragma_with_expansion functions to register custom pragmas.

```c
/* Plugin callback called during pragmas registration. Registered with
   register_callback (plugin_name, PLUGIN_PRAGMAS,
   register_my_pragma, NULL);
*/
static void
register_my_pragma (void *event_data, void *data)
{
  warning (0, G_("Callback to register pragmas"));
  c_register_pragma ("GCCPLUGIN", "sayhello", handle_pragma_sayhello);
}
```

It is suggested to pass "GCCPLUGIN" (or a short name identifying your plugin) as the “space” argument of your pragma.

### 23.7 Recording information about pass execution

The event PLUGIN_PASS_EXECUTION passes the pointer to the executed pass (the same as current_pass) as gcc_data to the callback. You can also inspect cfun to find out about which function this pass is executed for. Note that this event will only be invoked if the gate check (if applicable, modified by PLUGIN_OVERRIDE_GATE) succeeds. You can use other hooks, like PLUGIN_ALL_PASSES_START, PLUGIN_ALL_PASSES_END, PLUGIN_ALL_IPA_PASSES_START, PLUGIN_ALL_IPA_PASSES_END, PLUGIN_EARLY_GIMPLE_PASSES_START, and/or PLUGIN_EARLY_GIMPLE_PASSES_END to manipulate global state in your plugin(s) in order to get context for the pass execution.

### 23.8 Controlling which passes are being run

After the original gate function for a pass is called, its result - the gate status - is stored as an integer. Then the event PLUGIN_OVERRIDE_GATE is invoked, with a pointer to the gate status in the gcc_data parameter to the callback function. A nonzero value of the gate status means that the pass is to be executed. You can both read and write the gate status via the passed pointer.

### 23.9 Keeping track of available passes

When your plugin is loaded, you can inspect the various pass lists to determine what passes are available. However, other plugins might add new passes. Also, future changes to GCC might cause generic passes to be added after plugin loading. When a pass is first added to one of the pass lists, the event PLUGIN_NEW_PASS is invoked, with the callback parameter gcc_data pointing to the new pass.

### 23.10 Building GCC plugins

If plugins are enabled, GCC installs the headers needed to build a plugin (somewhere in the installation tree, e.g. under ‘/usr/local’). In particular a ‘plugin/include’ directory is installed, containing all the header files needed to build plugins.
On most systems, you can query this plugin directory by invoking gcc -print-file-name=plugin (replace if needed gcc with the appropriate program path).

Inside plugins, this plugin directory name can be queried by calling default_plugin_dir_name().

The following GNU Makefile excerpt shows how to build a simple plugin:

```makefile
GCC=gcc
PLUGIN_SOURCE_FILES= plugin1.c plugin2.c
PLUGIN_OBJECT_FILES= $(patsubst %.c,%.o,\$(PLUGIN_SOURCE_FILES))
GCCPLUGINS_DIR:= $(shell $(GCC) -print-file-name=plugin)
CFLAGS+= -I$(GCCPLUGINS_DIR)/include -fPIC -O2

plugin.so: $(PLUGIN_OBJECT_FILES)
    $(GCC) -shared $^ -o $@
```

A single source file plugin may be built with gcc -I'gcc -print-file-name=plugin'/include -fPIC -shared -O2 plugin.c -o plugin.so, using backquote shell syntax to query the ‘plugin’ directory.

Plugins needing to use gengtype require a GCC build directory for the same version of GCC that they will be linked against.
24 Link Time Optimization

24.1 Design Overview

Link time optimization is implemented as a GCC front end for a bytecode representation of GIMPLE that is emitted in special sections of .o files. Currently, LTO support is enabled in most ELF-based systems, as well as darwin, cygwin and mingw systems.

Since GIMPLE bytecode is saved alongside final object code, object files generated with LTO support are larger than regular object files. This “fat” object format makes it easy to integrate LTO into existing build systems, as one can, for instance, produce archives of the files. Additionally, one might be able to ship one set of fat objects which could be used both for development and the production of optimized builds. A, perhaps surprising, side effect of this feature is that any mistake in the toolchain that leads to LTO information not being used (e.g. an older libtool calling ld directly). This is both an advantage, as the system is more robust, and a disadvantage, as the user is not informed that the optimization has been disabled.

The current implementation only produces “fat” objects, effectively doubling compilation time and increasing file sizes up to 5x the original size. This hides the problem that some tools, such as ar and nm, need to understand symbol tables of LTO sections. These tools were extended to use the plugin infrastructure, and with these problems solved, GCC will also support “slim” objects consisting of the intermediate code alone.

At the highest level, LTO splits the compiler in two. The first half (the “writer”) produces a streaming representation of all the internal data structures needed to optimize and generate code. This includes declarations, types, the callgraph and the GIMPLE representation of function bodies.

When ‘-flto’ is given during compilation of a source file, the pass manager executes all the passes in all_lto_gen_passes. Currently, this phase is composed of two IPA passes:

- **pass_ipa_lto_gimple_out** This pass executes the function lto_output in ‘lto-streamer-out.c’, which traverses the call graph encoding every reachable declaration, type and function. This generates a memory representation of all the file sections described below.
- **pass_ipa_lto_finish_out** This pass executes the function produce_asm_for_decls in ‘lto-streamer-out.c’, which takes the memory image built in the previous pass and encodes it in the corresponding ELF file sections.

The second half of LTO support is the “reader”. This is implemented as the GCC front end ‘lto1’ in ‘lto/lto.c’. When ‘collect2’ detects a link set of .o/.a files with LTO information and the ‘-flto’ is enabled, it invokes ‘lto1’ which reads the set of files and aggregates them into a single translation unit for optimization. The main entry point for the reader is ‘lto/lto.c’:lto_main.

24.1.1 LTO modes of operation

One of the main goals of the GCC link-time infrastructure was to allow effective compilation of large programs. For this reason GCC implements two link-time compilation modes.

1. **LTO mode**, in which the whole program is read into the compiler at link-time and optimized in a similar way as if it were a single source-level compilation unit.
2. **WHOPR or partitioned mode**, designed to utilize multiple CPUs and/or a distributed compilation environment to quickly link large applications. WHOPR stands for WHOle Program optimizeR (not to be confused with the semantics of `--fwhole-program`). It partitions the aggregated callgraph from many different `.o` files and distributes the compilation of the sub-graphs to different CPUs.

   Note that distributed compilation is not implemented yet, but since the parallelism is facilitated via generating a Makefile, it would be easy to implement.

   WHOPR splits LTO into three main stages:

   1. Local generation (LGEN) This stage executes in parallel. Every file in the program is compiled into the intermediate language and packaged together with the local callgraph and summary information. This stage is the same for both the LTO and WHOPR compilation mode.

   2. Whole Program Analysis (WPA) WPA is performed sequentially. The global call-graph is generated, and a global analysis procedure makes transformation decisions. The global call-graph is partitioned to facilitate parallel optimization during phase 3. The results of the WPA stage are stored into new object files which contain the partitions of program expressed in the intermediate language and the optimization decisions.

   3. Local transformations (LTRANS) This stage executes in parallel. All the decisions made during phase 2 are implemented locally in each partitioned object file, and the final object code is generated. Optimizations which cannot be decided efficiently during the phase 2 may be performed on the local call-graph partitions.

   WHOPR can be seen as an extension of the usual LTO mode of compilation. In LTO, WPA and LTRANS are executed within a single execution of the compiler, after the whole program has been read into memory.

   When compiling in WHOPR mode, the callgraph is partitioned during the WPA stage. The whole program is split into a given number of partitions of roughly the same size. The compiler tries to minimize the number of references which cross partition boundaries. The main advantage of WHOPR is to allow the parallel execution of LTRANS stages, which are the most time-consuming part of the compilation process. Additionally, it avoids the need to load the whole program into memory.

### 24.2 LTO file sections

LTO information is stored in several ELF sections inside object files. Data structures and enum codes for sections are defined in `lto-streamer.h`.

These sections are emitted from `lto-streamer-out.c` and mapped in all at once from `lto/lto.c`: `lto_file_read`. The individual functions dealing with the reading/writing of each section are described below.

- **Command line options (.gnu.lto_.opts)**

  This section contains the command line options used to generate the object files. This is used at link time to determine the optimization level and other settings when they are not explicitly specified at the linker command line.

  Currently, GCC does not support combining LTO object files compiled with different set of the command line options into a single binary. At link time, the options given on
the command line and the options saved on all the files in a link-time set are applied globally. No attempt is made at validating the combination of flags (other than the usual validation done by option processing). This is implemented in `lto/lto.c`:lto_read_all_file_options.

- Symbol table (`.gnu.lto_.symtab`)  
  This table replaces the ELF symbol table for functions and variables represented in the LTO IL. Symbols used and exported by the optimized assembly code of “fat” objects might not match the ones used and exported by the intermediate code. This table is necessary because the intermediate code is less optimized and thus requires a separate symbol table.

  Additionally, the binary code in the “fat” object will lack a call to a function, since the call was optimized out at compilation time after the intermediate language was streamed out. In some special cases, the same optimization may not happen during link-time optimization. This would lead to an undefined symbol if only one symbol table was used.

  The symbol table is emitted in `lto-streamer-out.c`:produce_symtab.

- Global declarations and types (`.gnu.lto_.decls`)  
  This section contains an intermediate language dump of all declarations and types required to represent the callgraph, static variables and top-level debug info.

  The contents of this section are emitted in `lto-streamer-out.c`:produce_asm_for_decls. Types and symbols are emitted in a topological order that preserves the sharing of pointers when the file is read back in (`lto.c`:read_cgraph_and_symbols).

- The callgraph (`.gnu.lto_.cgraph`)  
  This section contains the basic data structure used by the GCC inter-procedural optimization infrastructure. This section stores an annotated multi-graph which represents the functions and call sites as well as the variables, aliases and top-level `asm` statements.

  This section is emitted in `lto-streamer-out.c`:output_cgraph and read in `lto-cgraph.c`:input_cgraph.

- IPA references (`.gnu.lto_.refs`)  
  This section contains references between function and static variables. It is emitted by `lto-cgraph.c`:output.refs and read by `lto-cgraph.c`:input.refs.

- Function bodies (`.gnu.lto_.function_body.<name>`)  
  This section contains function bodies in the intermediate language representation. Every function body is in a separate section to allow copying of the section independently to different object files or reading the function on demand.

  Functions are emitted in `lto-streamer-out.c`:output_function and read in `lto-streamer-in.c`:input_function.

- Static variable initializers (`.gnu.lto_.vars`)  
  This section contains all the symbols in the global variable pool. It is emitted by `lto-cgraph.c`:output_varpool and read in `lto-cgraph.c`:input_cgraph.

- Summaries and optimization summaries used by IPA passes (`.gnu.lto_.<xxx>`, where `<xxx>` is one of jmpfuncs, pureconst or reference)
These sections are used by IPA passes that need to emit summary information during LTO generation to be read and aggregated at link time. Each pass is responsible for implementing two pass manager hooks: one for writing the summary and another for reading it in. The format of these sections is entirely up to each individual pass. The only requirement is that the writer and reader hooks agree on the format.

### 24.3 Using summary information in IPA passes

Programs are represented internally as a callgraph (a multi-graph where nodes are functions and edges are call sites) and a varpool (a list of static and external variables in the program).

The inter-procedural optimization is organized as a sequence of individual passes, which operate on the callgraph and the varpool. To make the implementation of WHOPR possible, every inter-procedural optimization pass is split into several stages that are executed at different times during WHOPR compilation:

- **LGEN time**
  1. *Generate summary* (*generate_summary* in *struct ipa_opt_pass_d*). This stage analyzes every function body and variable initializer is examined and stores relevant information into a pass-specific data structure.
  2. *Write summary* (*write_summary* in *struct ipa_opt_pass_d*). This stage writes all the pass-specific information generated by *generate_summary*. Summaries go into their own LTO_section_* sections that have to be declared in `lto-streamer.h`:enum lto_section_type. A new section is created by calling `create_output_block` and data can be written using the `lto_output_*` routines.

- **WPA time**
  1. *Read summary* (*read_summary* in *struct ipa_opt_pass_d*). This stage reads all the pass-specific information in exactly the same order that it was written by *write_summary*.
  2. *Execute* (*execute* in *struct opt_pass*). This performs inter-procedural propagation. This must be done without actual access to the individual function bodies or variable initializers. Typically, this results in a transitive closure operation over the summary information of all the nodes in the callgraph.
  3. *Write optimization summary* (*write_optimization_summary* in *struct ipa_opt_pass_d*). This writes the result of the inter-procedural propagation into the object file. This can use the same data structures and helper routines used in *write_summary*.

- **LTRANS time**
  1. *Read optimization summary* (*read_optimization_summary* in *struct ipa_opt_pass_d*). The counterpart to *write_optimization_summary*. This reads the interprocedural optimization decisions in exactly the same format emitted by *write_optimization_summary*.
  2. *Transform* (*function_transform* and *variable_transform* in *struct ipa_opt_pass_d*). The actual function bodies and variable initializers are updated based on the information passed down from the *Execute* stage.

The implementation of the inter-procedural passes are shared between LTO, WHOPR and classic non-LTO compilation.
During the traditional file-by-file mode every pass executes its own Generate summary, Execute, and Transform stages within the single execution context of the compiler.

In LTO compilation mode, every pass uses Generate summary and Write summary stages at compilation time, while the Read summary, Execute, and Transform stages are executed at link time.

In WHOPR mode all stages are used.

To simplify development, the GCC pass manager differentiates between normal inter-procedural passes and small inter-procedural passes. A small inter-procedural pass (SIMPLE_IPA_PASS) is a pass that does everything at once and thus it can not be executed during WPA in WHOPR mode. It defines only the Execute stage and during this stage it accesses and modifies the function bodies. Such passes are useful for optimization at LGEN or LTRANS time and are used, for example, to implement early optimization before writing object files. The simple inter-procedural passes can also be used for easier prototyping and development of a new inter-procedural pass.

**24.3.1 Virtual clones**

One of the main challenges of introducing the WHOPR compilation mode was addressing the interactions between optimization passes. In LTO compilation mode, the passes are executed in a sequence, each of which consists of analysis (or Generate summary), propagation (or Execute) and Transform stages. Once the work of one pass is finished, the next pass sees the updated program representation and can execute. This makes the individual passes dependent on each other.

In WHOPR mode all passes first execute their Generate summary stage. Then summary writing marks the end of the LGEN stage. At WPA time, the summaries are read back into memory and all passes run the Execute stage. Optimization summaries are streamed and sent to LTRANS, where all the passes execute the Transform stage.

Most optimization passes split naturally into analysis, propagation and transformation stages. But some do not. The main problem arises when one pass performs changes and the following pass gets confused by seeing different callgraphs between the Transform stage and the Generate summary or Execute stage. This means that the passes are required to communicate their decisions with each other.

To facilitate this communication, the GCC callgraph infrastructure implements virtual clones, a method of representing the changes performed by the optimization passes in the callgraph without needing to update function bodies.

A virtual clone in the callgraph is a function that has no associated body, just a description of how to create its body based on a different function (which itself may be a virtual clone).

The description of function modifications includes adjustments to the function’s signature (which allows, for example, removing or adding function arguments), substitutions to perform on the function body, and, for inlined functions, a pointer to the function that it will be inlined into.

It is also possible to redirect any edge of the callgraph from a function to its virtual clone. This implies updating of the call site to adjust for the new function signature.

Most of the transformations performed by inter-procedural optimizations can be represented via virtual clones. For instance, a constant propagation pass can produce a virtual
clone of the function which replaces one of its arguments by a constant. The inliner can represent its decisions by producing a clone of a function whose body will be later integrated into a given function.

Using virtual clones, the program can be easily updated during the Execute stage, solving most of pass interactions problems that would otherwise occur during Transform.

Virtual clones are later materialized in the LTRANS stage and turned into real functions. Passes executed after the virtual clone were introduced also perform their Transform stage on new functions, so for a pass there is no significant difference between operating on a real function or a virtual clone introduced before its Execute stage.

Optimization passes then work on virtual clones introduced before their Execute stage as if they were real functions. The only difference is that clones are not visible during the Generate Summary stage.

To keep function summaries updated, the callgraph interface allows an optimizer to register a callback that is called every time a new clone is introduced as well as when the actual function or variable is generated or when a function or variable is removed. These hooks are registered in the Generate summary stage and allow the pass to keep its information intact until the Execute stage. The same hooks can also be registered during the Execute stage to keep the optimization summaries updated for the Transform stage.

24.3.2 IPA references
GCC represents IPA references in the callgraph. For a function or variable A, the IPA reference is a list of all locations where the address of A is taken and, when A is a variable, a list of all direct stores and reads to/from A. References represent an oriented multi-graph on the union of nodes of the callgraph and the varpool. See ‘ipa-reference.c’:ipa_reference_write_optimization_summary and ‘ipa-reference.c’:ipa_reference_read_optimization_summary for details.

24.3.3 Jump functions
Suppose that an optimization pass sees a function A and it knows the values of (some of) its arguments. The jump function describes the value of a parameter of a given function call in function A based on this knowledge.

Jump functions are used by several optimizations, such as the inter-procedural constant propagation pass and the devirtualization pass. The inliner also uses jump functions to perform inlining of callbacks.

24.4 Whole program assumptions, linker plugin and symbol visibilities
Link-time optimization gives relatively minor benefits when used alone. The problem is that propagation of inter-procedural information does not work well across functions and variables that are called or referenced by other compilation units (such as from a dynamically linked library). We say that such functions are variables are externally visible.

To make the situation even more difficult, many applications organize themselves as a set of shared libraries, and the default ELF visibility rules allow one to overwrite any externally visible symbol with a different symbol at runtime. This basically disables any optimizations across such functions and variables, because the compiler cannot be sure that the function
body it is seeing is the same function body that will be used at runtime. Any function or variable not declared \texttt{static} in the sources degrades the quality of inter-procedural optimization.

To avoid this problem the compiler must assume that it sees the whole program when doing link-time optimization. Strictly speaking, the whole program is rarely visible even at link-time. Standard system libraries are usually linked dynamically or not provided with the link-time information. In GCC, the whole program option ('\texttt{-fwhole-program}’) asserts that every function and variable defined in the current compilation unit is static, except for function \texttt{main} (note: at link time, the current unit is the union of all objects compiled with LTO). Since some functions and variables need to be referenced externally, for example by another DSO or from an assembler file, GCC also provides the function and variable attribute \texttt{externally_visible} which can be used to disable the effect of ‘\texttt{-fwhole-program}’ on a specific symbol.

The whole program mode assumptions are slightly more complex in C++, where inline functions in headers are put into \texttt{COMDAT} sections. \texttt{COMDAT} function and variables can be defined by multiple object files and their bodies are unified at link-time and dynamic link-time. \texttt{COMDAT} functions are changed to local only when their address is not taken and thus un-sharing them with a library is not harmful. \texttt{COMDAT} variables always remain externally visible, however for readonly variables it is assumed that their initializers cannot be overwritten by a different value.

GCC provides the function and variable attribute \texttt{visibility} that can be used to specify the visibility of externally visible symbols (or alternatively an ‘\texttt{-fdefault-visibility}’ command line option). ELF defines the \texttt{default}, \texttt{protected}, \texttt{hidden} and \texttt{internal} visibilities.

The most commonly used is visibility is \texttt{hidden}. It specifies that the symbol cannot be referenced from outside of the current shared library. Unfortunately, this information cannot be used directly by the link-time optimization in the compiler since the whole shared library also might contain non-LTO objects and those are not visible to the compiler.

GCC solves this problem using linker plugins. A \texttt{linker plugin} is an interface to the linker that allows an external program to claim the ownership of a given object file. The linker then performs the linking procedure by querying the plugin about the symbol table of the claimed objects and once the linking decisions are complete, the plugin is allowed to provide the final object file before the actual linking is made. The linker plugin obtains the symbol resolution information which specifies which symbols provided by the claimed objects are bound from the rest of a binary being linked.

Currently, the linker plugin works only in combination with the Gold linker, but a GNU ld implementation is under development.

GCC is designed to be independent of the rest of the toolchain and aims to support linkers without plugin support. For this reason it does not use the linker plugin by default. Instead, the object files are examined by \texttt{collect2} before being passed to the linker and objects found to have LTO sections are passed to \texttt{lto1} first. This mode does not work for library archives. The decision on what object files from the archive are needed depends on the actual linking and thus GCC would have to implement the linker itself. The resolution information is missing too and thus GCC needs to make an educated guess based on ‘\texttt{-fwhole-program}’. 
Without the linker plugin GCC also assumes that symbols are declared hidden and not referred by non-LTO code by default.

### 24.5 Internal flags controlling `lto1`

The following flags are passed into `lto1` and are not meant to be used directly from the command line.

- `-fwpa` This option runs the serial part of the link-time optimizer performing the inter-procedural propagation (WPA mode). The compiler reads in summary information from all inputs and performs an analysis based on summary information only. It generates object files for subsequent runs of the link-time optimizer where individual object files are optimized using both summary information from the WPA mode and the actual function bodies. It then drives the LTRANS phase.

- `-fltrans` This option runs the link-time optimizer in the local-transformation (LTRANS) mode, which reads in output from a previous run of the LTO in WPA mode. In the LTRANS mode, LTO optimizes an object and produces the final assembly.

- `-fltrans-output-list=file` This option specifies a file to which the names of LTRANS output files are written. This option is only meaningful in conjunction with `'-fwpa'`. 

Funding Free Software

If you want to have more free software a few years from now, it makes sense for you to help encourage people to contribute funds for its development. The most effective approach known is to encourage commercial redistributors to donate.

Users of free software systems can boost the pace of development by encouraging for-a-fee distributors to donate part of their selling price to free software developers—the Free Software Foundation, and others.

The way to convince distributors to do this is to demand it and expect it from them. So when you compare distributors, judge them partly by how much they give to free software development. Show distributors they must compete to be the one who gives the most.

To make this approach work, you must insist on numbers that you can compare, such as, “We will donate ten dollars to the Frobnitz project for each disk sold.” Don’t be satisfied with a vague promise, such as “A portion of the profits are donated,” since it doesn’t give a basis for comparison.

Even a precise fraction “of the profits from this disk” is not very meaningful, since creative accounting and unrelated business decisions can greatly alter what fraction of the sales price counts as profit. If the price you pay is $50, ten percent of the profit is probably less than a dollar; it might be a few cents, or nothing at all.

Some redistributors do development work themselves. This is useful too; but to keep everyone honest, you need to inquire how much they do, and what kind. Some kinds of development make much more long-term difference than others. For example, maintaining a separate version of a program contributes very little; maintaining the standard version of a program for the whole community contributes much. Easy new ports contribute little, since someone else would surely do them; difficult ports such as adding a new CPU to the GNU Compiler Collection contribute more; major new features or packages contribute the most.

By establishing the idea that supporting further development is “the proper thing to do” when distributing free software for a fee, we can assure a steady flow of resources into making more free software.

Copyright © 1994 Free Software Foundation, Inc.

Verbatim copying and redistribution of this section is permitted without royalty; alteration is not permitted.
The GNU Project and GNU/Linux

The GNU Project was launched in 1984 to develop a complete Unix-like operating system which is free software: the GNU system. (GNU is a recursive acronym for “GNU’s Not Unix”; it is pronounced “guh-NEW”.) Variants of the GNU operating system, which use the kernel Linux, are now widely used; though these systems are often referred to as “Linux”, they are more accurately called GNU/Linux systems.

For more information, see:

http://www.gnu.org/
GNU General Public License

Version 3, 29 June 2007


Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

Preamble

The GNU General Public License is a free, copyleft license for software and other kinds of works.

The licenses for most software and other practical works are designed to take away your freedom to share and change the works. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change all versions of a program—to make sure it remains free software for all its users. We, the Free Software Foundation, use the GNU General Public License for most of our software; it applies also to any other work released this way by its authors. You can apply it to your programs, too.

When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for them if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs, and that you know you can do these things.

To protect your rights, we need to prevent others from denying you these rights or asking you to surrender the rights. Therefore, you have certain responsibilities if you distribute copies of the software, or if you modify it: responsibilities to respect the freedom of others.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must pass on to the recipients the same freedoms that you received. You must make sure that they, too, receive or can get the source code. And you must show them these terms so they know their rights.

Developers that use the GNU GPL protect your rights with two steps: (1) assert copyright on the software, and (2) offer you this License giving you legal permission to copy, distribute and/or modify it.

For the developers’ and authors’ protection, the GPL clearly explains that there is no warranty for this free software. For both users’ and authors’ sake, the GPL requires that modified versions be marked as changed, so that their problems will not be attributed erroneously to authors of previous versions.

Some devices are designed to deny users access to install or run modified versions of the software inside them, although the manufacturer can do so. This is fundamentally incompatible with the aim of protecting users’ freedom to change the software. The systematic pattern of such abuse occurs in the area of products for individuals to use, which is precisely where it is most unacceptable. Therefore, we have designed this version of the GPL to prohibit the practice for those products. If such problems arise substantially in other domains, we stand ready to extend this provision to those domains in future versions of the GPL, as needed to protect the freedom of users.
Finally, every program is threatened constantly by software patents. States should not allow patents to restrict development and use of software on general-purpose computers, but in those that do, we wish to avoid the special danger that patents applied to a free program could make it effectively proprietary. To prevent this, the GPL assures that patents cannot be used to render the program non-free.

The precise terms and conditions for copying, distribution and modification follow.

TERMS AND CONDITIONS

0. Definitions.

“This License” refers to version 3 of the GNU General Public License.

“Copyright” also means copyright-like laws that apply to other kinds of works, such as semiconductor masks.

“The Program” refers to any copyrightable work licensed under this License. Each licensee is addressed as “you”. “Licensees” and “recipients” may be individuals or organizations.

To “modify” a work means to copy from or adapt all or part of the work in a fashion requiring copyright permission, other than the making of an exact copy. The resulting work is called a “modified version” of the earlier work or a work “based on” the earlier work.

A “covered work” means either the unmodified Program or a work based on the Program.

To “propagate” a work means to do anything with it that, without permission, would make you directly or secondarily liable for infringement under applicable copyright law, except executing it on a computer or modifying a private copy. Propagation includes copying, distribution (with or without modification), making available to the public, and in some countries other activities as well.

To “convey” a work means any kind of propagation that enables other parties to make or receive copies. Mere interaction with a user through a computer network, with no transfer of a copy, is not conveying.

An interactive user interface displays “Appropriate Legal Notices” to the extent that it includes a convenient and prominently visible feature that (1) displays an appropriate copyright notice, and (2) tells the user that there is no warranty for the work (except to the extent that warranties are provided), that licensees may convey the work under this License, and how to view a copy of this License. If the interface presents a list of user commands or options, such as a menu, a prominent item in the list meets this criterion.


The “source code” for a work means the preferred form of the work for making modifications to it. “Object code” means any non-source form of a work.

A “Standard Interface” means an interface that either is an official standard defined by a recognized standards body, or, in the case of interfaces specified for a particular programming language, one that is widely used among developers working in that language.
The “System Libraries” of an executable work include anything, other than the work as a whole, that (a) is included in the normal form of packaging a Major Component, but which is not part of that Major Component, and (b) serves only to enable use of the work with that Major Component, or to implement a Standard Interface for which an implementation is available to the public in source code form. A “Major Component”, in this context, means a major essential component (kernel, window system, and so on) of the specific operating system (if any) on which the executable work runs, or a compiler used to produce the work, or an object code interpreter used to run it.

The “Corresponding Source” for a work in object code form means all the source code needed to generate, install, and (for an executable work) run the object code and to modify the work, including scripts to control those activities. However, it does not include the work’s System Libraries, or general-purpose tools or generally available free programs which are used unmodified in performing those activities but which are not part of the work. For example, Corresponding Source includes interface definition files associated with source files for the work, and the source code for shared libraries and dynamically linked subprograms that the work is specifically designed to require, such as by intimate data communication or control flow between those subprograms and other parts of the work.

The Corresponding Source need not include anything that users can regenerate automatically from other parts of the Corresponding Source.

The Corresponding Source for a work in source code form is that same work.

2. Basic Permissions.

All rights granted under this License are granted for the term of copyright on the Program, and are irrevocable provided the stated conditions are met. This License explicitly affirms your unlimited permission to run the unmodified Program. The output from running a covered work is covered by this License only if the output, given its content, constitutes a covered work. This License acknowledges your rights of fair use or other equivalent, as provided by copyright law.

You may make, run and propagate covered works that you do not convey, without conditions so long as your license otherwise remains in force. You may convey covered works to others for the sole purpose of having them make modifications exclusively for you, or provide you with facilities for running those works, provided that you comply with the terms of this License in conveying all material for which you do not control copyright. Those thus making or running the covered works for you must do so exclusively on your behalf, under your direction and control, on terms that prohibit them from making any copies of your copyrighted material outside their relationship with you.

Conveying under any other circumstances is permitted solely under the conditions stated below. Sublicensing is not allowed; section 10 makes it unnecessary.


No covered work shall be deemed part of an effective technological measure under any applicable law fulfilling obligations under article 11 of the WIPO copyright treaty adopted on 20 December 1996, or similar laws prohibiting or restricting circumvention of such measures.
When you convey a covered work, you waive any legal power to forbid circumvention of technological measures to the extent such circumvention is effected by exercising rights under this License with respect to the covered work, and you disclaim any intention to limit operation or modification of the work as a means of enforcing, against the work’s users, your or third parties’ legal rights to forbid circumvention of technological measures.


You may convey verbatim copies of the Program’s source code as you receive it, in any medium, provided that you conspicuously and appropriately publish on each copy an appropriate copyright notice; keep intact all notices stating that this License and any non-permissive terms added in accord with section 7 apply to the code; keep intact all notices of the absence of any warranty; and give all recipients a copy of this License along with the Program.

You may charge any price or no price for each copy that you convey, and you may offer support or warranty protection for a fee.

5. Conveying Modified Source Versions.

You may convey a work based on the Program, or the modifications to produce it from the Program, in the form of source code under the terms of section 4, provided that you also meet all of these conditions:

a. The work must carry prominent notices stating that you modified it, and giving a relevant date.

b. The work must carry prominent notices stating that it is released under this License and any conditions added under section 7. This requirement modifies the requirement in section 4 to “keep intact all notices”.

c. You must license the entire work, as a whole, under this License to anyone who comes into possession of a copy. This License will therefore apply, along with any applicable section 7 additional terms, to the whole of the work, and all its parts, regardless of how they are packaged. This License gives no permission to license the work in any other way, but it does not invalidate such permission if you have separately received it.

d. If the work has interactive user interfaces, each must display Appropriate Legal Notices; however, if the Program has interactive interfaces that do not display Appropriate Legal Notices, your work need not make them do so.

A compilation of a covered work with other separate and independent works, which are not by their nature extensions of the covered work, and which are not combined with it such as to form a larger program, in or on a volume of a storage or distribution medium, is called an “aggregate” if the compilation and its resulting copyright are not used to limit the access or legal rights of the compilation’s users beyond what the individual works permit. Inclusion of a covered work in an aggregate does not cause this License to apply to the other parts of the aggregate.

6. Conveying Non-Source Forms.

You may convey a covered work in object code form under the terms of sections 4 and 5, provided that you also convey the machine-readable Corresponding Source under the terms of this License, in one of these ways:
a. Convey the object code in, or embodied in, a physical product (including a physical distribution medium), accompanied by the Corresponding Source fixed on a durable physical medium customarily used for software interchange.

b. Convey the object code in, or embodied in, a physical product (including a physical distribution medium), accompanied by a written offer, valid for at least three years and valid for as long as you offer spare parts or customer support for that product model, to give anyone who possesses the object code either (1) a copy of the Corresponding Source for all the software in the product that is covered by this License, on a durable physical medium customarily used for software interchange, for a price no more than your reasonable cost of physically performing this conveying of source, or (2) access to copy the Corresponding Source from a network server at no charge.

c. Convey individual copies of the object code with a copy of the written offer to provide the Corresponding Source. This alternative is allowed only occasionally and noncommercially, and only if you received the object code with such an offer, in accord with subsection 6b.

d. Convey the object code by offering access from a designated place (gratis or for a charge), and offer equivalent access to the Corresponding Source in the same way through the same place at no further charge. You need not require recipients to copy the Corresponding Source along with the object code. If the place to copy the object code is a network server, the Corresponding Source may be on a different server (operated by you or a third party) that supports equivalent copying facilities, provided you maintain clear directions next to the object code saying where to find the Corresponding Source. Regardless of what server hosts the Corresponding Source, you remain obligated to ensure that it is available for as long as needed to satisfy these requirements.

e. Convey the object code using peer-to-peer transmission, provided you inform other peers where the object code and Corresponding Source of the work are being offered to the general public at no charge under subsection 6d.

A separable portion of the object code, whose source code is excluded from the Corresponding Source as a System Library, need not be included in conveying the object code work.

A “User Product” is either (1) a “consumer product”, which means any tangible personal property which is normally used for personal, family, or household purposes, or (2) anything designed or sold for incorporation into a dwelling. In determining whether a product is a consumer product, doubtful cases shall be resolved in favor of coverage. For a particular product received by a particular user, “normally used” refers to a typical or common use of that class of product, regardless of the status of the particular user or of the way in which the particular user actually uses, expects or is expected to use, the product. A product is a consumer product regardless of whether the product has substantial commercial, industrial or non-consumer uses, unless such uses represent the only significant mode of use of the product.

“Installation Information” for a User Product means any methods, procedures, authorization keys, or other information required to install and execute modified versions of a covered work in that User Product from a modified version of its Corresponding Source.
The information must suffice to ensure that the continued functioning of the modified object code is in no case prevented or interfered with solely because modification has been made.

If you convey an object code work under this section in, or with, or specifically for use in, a User Product, and the conveying occurs as part of a transaction in which the right of possession and use of the User Product is transferred to the recipient in perpetuity or for a fixed term (regardless of how the transaction is characterized), the Corresponding Source conveyed under this section must be accompanied by the Installation Information. But this requirement does not apply if neither you nor any third party retains the ability to install modified object code on the User Product (for example, the work has been installed in ROM).

The requirement to provide Installation Information does not include a requirement to continue to provide support service, warranty, or updates for a work that has been modified or installed by the recipient, or for the User Product in which it has been modified or installed. Access to a network may be denied when the modification itself materially and adversely affects the operation of the network or violates the rules and protocols for communication across the network.

Corresponding Source conveyed, and Installation Information provided, in accord with this section must be in a format that is publicly documented (and with an implementation available to the public in source code form), and must require no special password or key for unpacking, reading or copying.

7. Additional Terms.

"Additional permissions" are terms that supplement the terms of this License by making exceptions from one or more of its conditions. Additional permissions that are applicable to the entire Program shall be treated as though they were included in this License, to the extent that they are valid under applicable law. If additional permissions apply only to part of the Program, that part may be used separately under those permissions, but the entire Program remains governed by this License without regard to the additional permissions.

When you convey a copy of a covered work, you may at your option remove any additional permissions from that copy, or from any part of it. (Additional permissions may be written to require their own removal in certain cases when you modify the work.) You may place additional permissions on material, added by you to a covered work, for which you have or can give appropriate copyright permission.

Notwithstanding any other provision of this License, for material you add to a covered work, you may (if authorized by the copyright holders of that material) supplement the terms of this License with terms:

a. Disclaiming warranty or limiting liability differently from the terms of sections 15 and 16 of this License; or

b. Requiring preservation of specified reasonable legal notices or author attributions in that material or in the Appropriate Legal Notices displayed by works containing it; or

c. Prohibiting misrepresentation of the origin of that material, or requiring that modified versions of such material be marked in reasonable ways as different from the original version; or
d. Limiting the use for publicity purposes of names of licensors or authors of the material; or

e. Declining to grant rights under trademark law for use of some trade names, trademarks, or service marks; or

f. Requiring indemnification of licensors and authors of that material by anyone who conveys the material (or modified versions of it) with contractual assumptions of liability to the recipient, for any liability that these contractual assumptions directly impose on those licensors and authors.

All other non-permissive additional terms are considered “further restrictions” within the meaning of section 10. If the Program as you received it, or any part of it, contains a notice stating that it is governed by this License along with a term that is a further restriction, you may remove that term. If a license document contains a further restriction but permits relicensing or conveying under this License, you may add to a covered work material governed by the terms of that license document, provided that the further restriction does not survive such relicensing or conveying.

If you add terms to a covered work in accord with this section, you must place, in the relevant source files, a statement of the additional terms that apply to those files, or a notice indicating where to find the applicable terms.

Additional terms, permissive or non-permissive, may be stated in the form of a separately written license, or stated as exceptions; the above requirements apply either way.

8. Termination.

You may not propagate or modify a covered work except as expressly provided under this License. Any attempt otherwise to propagate or modify it is void, and will automatically terminate your rights under this License (including any patent licenses granted under the third paragraph of section 11).

However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, you do not qualify to receive new licenses for the same material under section 10.


You are not required to accept this License in order to receive or run a copy of the Program. Ancillary propagation of a covered work occurring solely as a consequence of using peer-to-peer transmission to receive a copy likewise does not require acceptance.
However, nothing other than this License grants you permission to propagate or modify any covered work. These actions infringe copyright if you do not accept this License. Therefore, by modifying or propagating a covered work, you indicate your acceptance of this License to do so.

10. Automatic Licensing of Downstream Recipients.

Each time you convey a covered work, the recipient automatically receives a license from the original licensors, to run, modify and propagate that work, subject to this License. You are not responsible for enforcing compliance by third parties with this License.

An “entity transaction” is a transaction transferring control of an organization, or substantially all assets of one, or subdividing an organization, or merging organizations. If propagation of a covered work results from an entity transaction, each party to that transaction who receives a copy of the work also receives whatever licenses to the work the party’s predecessor in interest had or could give under the previous paragraph, plus a right to possession of the Corresponding Source of the work from the predecessor in interest, if the predecessor has it or can get it with reasonable efforts.

You may not impose any further restrictions on the exercise of the rights granted or affirmed under this License. For example, you may not impose a license fee, royalty, or other charge for exercise of rights granted under this License, and you may not initiate litigation (including a cross-claim or counterclaim in a lawsuit) alleging that any patent claim is infringed by making, using, selling, offering for sale, or importing the Program or any portion of it.

11. Patents.

A “contributor” is a copyright holder who authorizes use under this License of the Program or a work on which the Program is based. The work thus licensed is called the contributor’s “contributor version”.

A contributor’s “essential patent claims” are all patent claims owned or controlled by the contributor, whether already acquired or hereafter acquired, that would be infringed by some manner, permitted by this License, of making, using, or selling its contributor version, but do not include claims that would be infringed only as a consequence of further modification of the contributor version. For purposes of this definition, “control” includes the right to grant patent sublicenses in a manner consistent with the requirements of this License.

Each contributor grants you a non-exclusive, worldwide, royalty-free patent license under the contributor’s essential patent claims, to make, use, sell, offer for sale, import and otherwise run, modify and propagate the contents of its contributor version.

In the following three paragraphs, a “patent license” is any express agreement or commitment, however denominated, not to enforce a patent (such as an express permission to practice a patent or covenant not to sue for patent infringement). To “grant” such a patent license to a party means to make such an agreement or commitment not to enforce a patent against the party.

If you convey a covered work, knowingly relying on a patent license, and the Corresponding Source of the work is not available for anyone to copy, free of charge and under the terms of this License, through a publicly available network server or other readily accessible means, then you must either (1) cause the Corresponding Source to be so
available, or (2) arrange to deprive yourself of the benefit of the patent license for this particular work, or (3) arrange, in a manner consistent with the requirements of this License, to extend the patent license to downstream recipients. “Knowingly relying” means you have actual knowledge that, but for the patent license, your conveying the covered work in a country, or your recipient’s use of the covered work in a country, would infringe one or more identifiable patents in that country that you have reason to believe are valid.

If, pursuant to or in connection with a single transaction or arrangement, you convey, or propagate by procuring conveyance of, a covered work, and grant a patent license to some of the parties receiving the covered work authorizing them to use, propagate, modify or convey a specific copy of the covered work, then the patent license you grant is automatically extended to all recipients of the covered work and works based on it.

A patent license is “discriminatory” if it does not include within the scope of its coverage, prohibits the exercise of, or is conditioned on the non-exercise of one or more of the rights that are specifically granted under this License. You may not convey a covered work if you are a party to an arrangement with a third party that is in the business of distributing software, under which you make payment to the third party based on the extent of your activity of conveying the work, and under which the third party grants, to any of the parties who would receive the covered work from you, a discriminatory patent license (a) in connection with copies of the covered work conveyed by you (or copies made from those copies), or (b) primarily for and in connection with specific products or compilations that contain the covered work, unless you entered into that arrangement, or that patent license was granted, prior to 28 March 2007.

Nothing in this License shall be construed as excluding or limiting any implied license or other defenses to infringement that may otherwise be available to you under applicable patent law.


If conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot convey a covered work so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not convey it at all. For example, if you agree to terms that obligate you to collect a royalty for further conveying from those to whom you convey the Program, the only way you could satisfy both those terms and this License would be to refrain entirely from conveying the Program.

13. Use with the GNU Affero General Public License.

Notwithstanding any other provision of this License, you have permission to link or combine any covered work with a work licensed under version 3 of the GNU Affero General Public License into a single combined work, and to convey the resulting work. The terms of this License will continue to apply to the part which is the covered work, but the special requirements of the GNU Affero General Public License, section 13, concerning interaction through a network will apply to the combination as such.

14. Revised Versions of this License.
The Free Software Foundation may publish revised and/or new versions of the GNU General Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. Each version is given a distinguishing version number. If the Program specifies that a certain numbered version of the GNU General Public License “or any later version” applies to it, you have the option of following the terms and conditions either of that numbered version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of the GNU General Public License, you may choose any version ever published by the Free Software Foundation.

If the Program specifies that a proxy can decide which future versions of the GNU General Public License can be used, that proxy’s public statement of acceptance of a version permanently authorizes you to choose that version for the Program.

Later license versions may give you additional or different permissions. However, no additional obligations are imposed on any author or copyright holder as a result of your choosing to follow a later version.

15. Disclaimer of Warranty.

THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

16. Limitation of Liability.

IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MODIFIES AND/OR CONVEYS THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

17. Interpretation of Sections 15 and 16.

If the disclaimer of warranty and limitation of liability provided above cannot be given local legal effect according to their terms, reviewing courts shall apply local law that most closely approximates an absolute waiver of all civil liability in connection with the Program, unless a warranty or assumption of liability accompanies a copy of the Program in return for a fee.
END OF TERMS AND CONDITIONS

How to Apply These Terms to Your New Programs

If you develop a new program, and you want it to be of the greatest possible use to the public, the best way to achieve this is to make it free software which everyone can redistribute and change under these terms.

To do so, attach the following notices to the program. It is safest to attach them to the start of each source file to most effectively state the exclusion of warranty; and each file should have at least the “copyright” line and a pointer to where the full notice is found.

one line to give the program’s name and a brief idea of what it does.
Copyright (C) year name of author

This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program. If not, see http://www.gnu.org/licenses/.

Also add information on how to contact you by electronic and paper mail.

If the program does terminal interaction, make it output a short notice like this when it starts in an interactive mode:

program Copyright (C) year name of author
This program comes with ABSOLUTELY NO WARRANTY; for details type ‘show w’.
This is free software, and you are welcome to redistribute it under certain conditions; type ‘show c’ for details.

The hypothetical commands ‘show w’ and ‘show c’ should show the appropriate parts of the General Public License. Of course, your program’s commands might be different; for a GUI interface, you would use an “about box”.

You should also get your employer (if you work as a programmer) or school, if any, to sign a “copyright disclaimer” for the program, if necessary. For more information on this, and how to apply and follow the GNU GPL, see http://www.gnu.org/licenses/.

The GNU General Public License does not permit incorporating your program into proprietary programs. If your program is a subroutine library, you may consider it more useful to permit linking proprietary applications with the library. If this is what you want to do, use the GNU Lesser General Public License instead of this License. But first, please read http://www.gnu.org/philosophy/why-not-lgpl.html.
GNU Free Documentation License

Version 1.3, 3 November 2008


http://fsf.org/

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

0. PREAMBLE
The purpose of this License is to make a manual, textbook, or other functional and useful document free in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or non-commercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of “copyleft”, which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS
This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The “Document”, below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as “you”. You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A “Modified Version” of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A “Secondary Section” is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document’s overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The “Invariant Sections” are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released...
under this License. If a section does not fit the above definition of Secondary then it is
not allowed to be designated as Invariant. The Document may contain zero Invariant
Sections. If the Document does not identify any Invariant Sections then there are none.

The “Cover Texts” are certain short passages of text that are listed, as Front-Cover
Texts or Back-Cover Texts, in the notice that says that the Document is released under
this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may
be at most 25 words.

A “Transparent” copy of the Document means a machine-readable copy, represented
in a format whose specification is available to the general public, that is suitable for
revising the document straightforwardly with generic text editors or (for images com-
posed of pixels) generic paint programs or (for drawings) some widely available drawing
editor, and that is suitable for input to text formatters or for automatic translation to
a variety of formats suitable for input to text formatters. A copy made in an otherwise
Transparent file format whose markup, or absence of markup, has been arranged to
thwart or discourage subsequent modification by readers is not Transparent. An image
format is not Transparent if used for any substantial amount of text. A copy that is
not “Transparent” is called “Opaque”.

Examples of suitable formats for Transparent copies include plain ASCII without
markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly
available DTD, and standard-conforming simple HTML, PostScript or PDF designed
for human modification. Examples of transparent image formats include PNG, XCF
and JPG. Opaque formats include proprietary formats that can be read and edited
only by proprietary word processors, SGML or XML for which the DTD and/or
processing tools are not generally available, and the machine-generated HTML,
PostScript or PDF produced by some word processors for output purposes only.

The “Title Page” means, for a printed book, the title page itself, plus such following
pages as are needed to hold, legibly, the material this License requires to appear in the
title page. For works in formats which do not have any title page as such, “Title Page”
means the text near the most prominent appearance of the work’s title, preceding the
beginning of the body of the text.

The “publisher” means any person or entity that distributes copies of the Document
to the public.

A section “Entitled XYZ” means a named subunit of the Document whose title either
is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in
another language. (Here XYZ stands for a specific section name mentioned below, such
as “Acknowledgements”, “Dedications”, “Endorsements”, or “History”.) To “Preserve
the Title” of such a section when you modify the Document means that it remains a
section “Entitled XYZ” according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that
this License applies to the Document. These Warranty Disclaimers are considered to
be included by reference in this License, but only as regards disclaiming warranties:
any other implication that these Warranty Disclaimers may have is void and has no
effect on the meaning of this License.

2. VERBATIM COPYING
You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document’s license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any,
be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled “History”, Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled “History” in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the “History” section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled “Acknowledgements” or “Dedications”, Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled “Endorsements”. Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled “Endorsements” or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their
titles to the list of Invariant Sections in the Modified Version’s license notice. These
titles must be distinct from any other section titles.

You may add a section Entitled “Endorsements”, provided it contains nothing but
endorsements of your Modified Version by various parties—for example, statements of
peer review or that the text has been approved by an organization as the authoritative
definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up
to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified
Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be
added by (or through arrangements made by) any one entity. If the Document already
includes a cover text for the same cover, previously added by you or by arrangement
made by the same entity you are acting on behalf of, you may not add another; but
you may replace the old one, on explicit permission from the previous publisher that
added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission
to use their names for publicity for or to assert or imply endorsement of any Modified
Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License,
under the terms defined in section 4 above for modified versions, provided that you
include in the combination all of the Invariant Sections of all of the original documents,
unmodified, and list them all as Invariant Sections of your combined work in its license
notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical
Invariant Sections may be replaced with a single copy. If there are multiple Invariant
Sections with the same name but different contents, make the title of each such section
unique by adding at the end of it, in parentheses, the name of the original author or
publisher of that section if known, or else a unique number. Make the same adjustment
to the section titles in the list of Invariant Sections in the license notice of the combined
work.

In the combination, you must combine any sections Entitled “History” in the vari-
ous original documents, forming one section Entitled “History”; likewise combine any
sections Entitled “Acknowledgements”, and any sections Entitled “Dedications”. You
must delete all sections Entitled “Endorsements.”

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released
under this License, and replace the individual copies of this License in the various
documents with a single copy that is included in the collection, provided that you
follow the rules of this License for verbatim copying of each of the documents in all
other respects.

You may extract a single document from such a collection, and distribute it individu-
ally under this License, provided you insert a copy of this License into the extracted
document, and follow this License in all other respects regarding verbatim copying of
that document.
7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent
documents or works, in or on a volume of a storage or distribution medium, is called
an “aggregate” if the copyright resulting from the compilation is not used to limit the
legal rights of the compilation’s users beyond what the individual works permit. When
the Document is included in an aggregate, this License does not apply to the other
works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document,
then if the Document is less than one half of the entire aggregate, the Document’s Cover
Texts may be placed on covers that bracket the Document within the aggregate, or the
electronic equivalent of covers if the Document is in electronic form. Otherwise they
must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations
of the Document under the terms of section 4. Replacing Invariant Sections with
translations requires special permission from their copyright holders, but you may
include translations of some or all Invariant Sections in addition to the original versions
of these Invariant Sections. You may include a translation of this License, and all the
license notices in the Document, and any Warranty Disclaimers, provided that you
also include the original English version of this License and the original versions of
those notices and disclaimers. In case of a disagreement between the translation and
the original version of this License or a notice or disclaimer, the original version will
prevail.

If a section in the Document is Entitled “Acknowledgements”, “Dedications”, or “His-
tory”, the requirement (section 4) to Preserve its Title (section 1) will typically require
changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly
provided under this License. Any attempt otherwise to copy, modify, sublicense, or
distribute it is void, and will automatically terminate your rights under this License.

However, if you cease all violation of this License, then your license from a particular
copyright holder is reinstated (a) provisionally, unless and until the copyright holder
explicitly and finally terminates your license, and (b) permanently, if the copyright
holder fails to notify you of the violation by some reasonable means prior to 60 days
after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if
the copyright holder notifies you of the violation by some reasonable means, this is the
first time you have received notice of violation of this License (for any work) from that
copyright holder, and you cure the violation prior to 30 days after your receipt of the
notice.

Termination of your rights under this section does not terminate the licenses of parties
who have received copies or rights from you under this License. If your rights have
been terminated and not permanently reinstated, receipt of a copy of some or all of the
same material does not give you any rights to use it.
10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License “or any later version” applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation. If the Document specifies that a proxy can decide which future versions of this License can be used, that proxy’s public statement of acceptance of a version permanently authorizes you to choose that version for the Document.

11. RELICENSING

“Massive Multiauthor Collaboration Site” (or “MMC Site”) means any World Wide Web server that publishes copyrightable works and also provides prominent facilities for anybody to edit those works. A public wiki that anybody can edit is an example of such a server. A “Massive Multiauthor Collaboration” (or “MMC”) contained in the site means any set of copyrightable works thus published on the MMC site.

“CC-BY-SA” means the Creative Commons Attribution-Share Alike 3.0 license published by Creative Commons Corporation, a not-for-profit corporation with a principal place of business in San Francisco, California, as well as future copyleft versions of that license published by that same organization.

“Incorporate” means to publish or republish a Document, in whole or in part, as part of another Document.

An MMC is “eligible for relicensing” if it is licensed under this License, and if all works that were first published under this License somewhere other than this MMC, and subsequently incorporated in whole or in part into the MMC, (1) had no cover texts or invariant sections, and (2) were thus incorporated prior to November 1, 2008.

The operator of an MMC Site may republish an MMC contained in the site under CC-BY-SA on the same site at any time before August 1, 2009, provided the MMC is eligible for relicensing.
ADDENDUM: How to use this License for your documents

To use this License in a document you have written, include a copy of the License in the document and put the following copyright and license notices just after the title page:

Copyright (C) year your name.
Permission is granted to copy, distribute and/or modify this document
under the terms of the GNU Free Documentation License, Version 1.3
or any later version published by the Free Software Foundation;
with no Invariant Sections, no Front-Cover Texts, and no Back-Cover
Texts. A copy of the license is included in the section entitled "GNU
Free Documentation License".

If you have Invariant Sections, Front-Cover Texts and Back-Cover Texts, replace the "with...Texts." line with this:

with the Invariant Sections being list their titles, with
the Front-Cover Texts being list, and with the Back-Cover Texts
being list.

If you have Invariant Sections without Cover Texts, or some other combination of the three, merge those two alternatives to suit the situation.

If your document contains nontrivial examples of program code, we recommend releasing these examples in parallel under your choice of free software license, such as the GNU General Public License, to permit their use in free software.
Contributors to GCC

The GCC project would like to thank its many contributors. Without them the project would not have been nearly as successful as it has been. Any omissions in this list are accidental. Feel free to contact law@redhat.com or gerald@pfeifer.com if you have been left out or some of your contributions are not listed. Please keep this list in alphabetical order.

- Analog Devices helped implement the support for complex data types and iterators.
- John David Anglin for threading-related fixes and improvements to libstdc++-v3, and the HP-UX port.
- James van Artsdalen wrote the code that makes efficient use of the Intel 80387 register stack.
- Abramo and Roberto Bagnara for the SysV68 Motorola 3300 Delta Series port.
- Alasdair Baird for various bug fixes.
- Giovanni Bajo for analyzing lots of complicated C++ problem reports.
- Peter Barada for his work to improve code generation for new ColdFire cores.
- Gerald Baumgartner added the signature extension to the C++ front end.
- Godmar Back for his Java improvements and encouragement.
- Scott Bambrough for help porting the Java compiler.
- Wolfgang Bangerth for processing tons of bug reports.
- Jon Beniston for his Microsoft Windows port of Java and port to Lattice Mico32.
- Daniel Berlin for better DWARF2 support, faster/better optimizations, improved alias analysis, plus migrating GCC to Bugzilla.
- Geoff Berry for his Java object serialization work and various patches.
- Uros Bizjak for the implementation of x87 math built-in functions and for various middle end and i386 back end improvements and bug fixes.
- Eric Blake for helping to make GCJ and libgcj conform to the specifications.
- Janne Blomqvist for contributions to GNU Fortran.
- Segher Boessenkool for various fixes.
- Hans-J. Boehm for his garbage collector, IA-64 libffi port, and other Java work.
- Neil Booth for work on cpplib, lang hooks, debug hooks and other miscellaneous clean-ups.
- Steven Bosscher for integrating the GNU Fortran front end into GCC and for contributing to the tree-ssa branch.
- Eric Botcazou for fixing middle- and backend bugs left and right.
- Per Bothner for his direction via the steering committee and various improvements to the infrastructure for supporting new languages. Chill front end implementation. Initial implementations of cpplib, fix-header, config.guess, libio, and past C++ library (libg++) maintainer. Dreaming up, designing and implementing much of GCJ.
- Devon Bowen helped port GCC to the Tahoe.
- Don Bowman for mips-vxworks contributions.
• Dave Brolley for work on cpplib and Chill.
• Paul Brook for work on the ARM architecture and maintaining GNU Fortran.
• Robert Brown implemented the support for Encore 32000 systems.
• Christian Bruel for improvements to local store elimination.
• Herman A.J. ten Brugge for various fixes.
• Joerg Brunsmann for Java compiler hacking and help with the GCJ FAQ.
• Joe Buck for his direction via the steering committee.
• Craig Burley for leadership of the G77 Fortran effort.
• Stephan Buys for contributing Doxygen notes for libstdc++.
• Paolo Carlini for libstdc++ work: lots of efficiency improvements to the C++ strings, streambufs and formatted I/O, hard detective work on the frustrating localization issues, and keeping up with the problem reports.
• John Carr for his alias work, SPARC hacking, infrastructure improvements, previous contributions to the steering committee, loop optimizations, etc.
• Stephane Carrez for 68HC11 and 68HC12 ports.
• Steve Chamberlain for support for the Renesas SH and H8 processors and the PicoJava processor, and for GCJ config fixes.
• Glenn Chambers for help with the GCJ FAQ.
• John-Marc Chandonia for various libgcj patches.
• Denis Chertykov for contributing and maintaining the AVR port, the first GCC port for an 8-bit architecture.
• Scott Christley for his Objective-C contributions.
• Eric Christopher for his Java porting help and clean-ups.
• Branko Cibej for more warning contributions.
• The GNU Classpath project for all of their merged runtime code.
• Nick Clifton for arm, mcore, fr30, v850, m32r, rx work, ‘--help’, and other random hacking.
• Michael Cook for libstdc++ cleanup patches to reduce warnings.
• R. Kelley Cook for making GCC buildable from a read-only directory as well as other miscellaneous build process and documentation clean-ups.
• Ralf Corsepius for SH testing and minor bug fixing.
• Stan Cox for care and feeding of the x86 port and lots of behind the scenes hacking.
• Alex Crain provided changes for the 3b1.
• Ian Dall for major improvements to the NS32k port.
• Paul Dale for his work to add uClinux platform support to the m68k backend.
• Dario Dariol contributed the four varieties of sample programs that print a copy of their source.
• Russell Davidson for fstream and stringstream fixes in libstdc++.
• Bud Davis for work on the G77 and GNU Fortran compilers.
• Mo DeJong for GCJ and libgcj bug fixes.
- DJ Delorie for the DJGPP port, build and libiberty maintenance, various bug fixes, and the M32C and MeP ports.
- Arnaud Desitter for helping to debug GNU Fortran.
- Gabriel Dos Reis for contributions to G++, contributions and maintenance of GCC diagnostics infrastructure, libstdc++-v3, including `valarray<>`, `complex<>`, maintaining the numerics library (including that pesky `<limits>` :-) and keeping up-to-date anything to do with numbers.
- Ulrich Drepper for his work on glibc, testing of GCC using glibc, ISO C99 support, CFG dumping support, etc., plus support of the C++ runtime libraries including for all kinds of C interface issues, contributing and maintaining `complex<>`, sanity checking and disbursement, configuration architecture, libio maintenance, and early math work.
- Zdenek Dvorak for a new loop unroller and various fixes.
- Michael Eager for his work on the Xilinix MicroBlaze port.
- Richard Earnshaw for his ongoing work with the ARM.
- David Edelsohn for his direction via the steering committee, ongoing work with the RS6000/PowerPC port, help cleaning up Haifa loop changes, doing the entire AIX port of libstdc++ with his bare hands, and for ensuring GCC properly keeps working on AIX.
- Kevin Ediger for the floating point formatting of `num_put::do_put` in libstdc++.
- Phil Edwards for libstdc++ work including configuration hackery, documentation maintainer, chief breaker of the web pages, the occasional istream bug fix, and work on shared library symbol versioning.
- Paul Eggert for random hacking all over GCC.
- Mark Elbrecht for various DJGPP improvements, and for libstdc++ configuration support for locales and fstream-related fixes.
- Vadim Egorov for libstdc++ fixes in strings, streambufs, and istreams.
- Christian Ehrhardt for dealing with bug reports.
- Ben Elliston for his work to move the Objective-C runtime into its own subdirectory and for his work on autoconf.
- Revital Eres for work on the PowerPC 750CL port.
- Marc Espie for OpenBSD support.
- Doug Evans for much of the global optimization framework, arc, m32r, and SPARC work.
- Christopher Faylor for his work on the Cygwin port and for caring and feeding the gcc.gnu.org box and saving its users tons of spam.
- Fred Fish for BeOS support and Ada fixes.
- Ivan Fontes Garcia for the Portuguese translation of the GCJ FAQ.
- Peter Gerwinski for various bug fixes and the Pascal front end.
- Kaveh R. Ghazi for his direction via the steering committee, amazing work to make `-W -Wall -W* -Werror` useful, and continuously testing GCC on a plethora of platforms. Kaveh extends his gratitude to the CAIP Center at Rutgers University for providing him with computing resources to work on Free Software since the late 1980s.
• John Gilmore for a donation to the FSF earmarked improving GNU Java.
• Judy Goldberg for c++ contributions.
• Torbjorn Granlund for various fixes and the c-torture testsuite, multiply- and divide-by-constant optimization, improved long long support, improved leaf function register allocation, and his direction via the steering committee.
• Anthony Green for his ‘-Os’ contributions, the moxie port, and Java front end work.
• Stu Grossman for gdb hacking, allowing GCJ developers to debug Java code.
• Michael K. Gschwind contributed the port to the PDP-11.
• Richard Guenther for his ongoing middle-end contributions and bug fixes and for release management.
• Ron Guilmette implemented the protoize and unprotoize tools, the support for Dwarf symbolic debugging information, and much of the support for System V Release 4. He has also worked heavily on the Intel 386 and 860 support.
• Mostafa Hagog for Swing Modulo Scheduling (SMS) and post reload GCSE.
• Bruno Haible for improvements in the runtime overhead for EH, new warnings and assorted bug fixes.
• Andrew Haley for his amazing Java compiler and library efforts.
• Chris Hanson assisted in making GCC work on HP-UX for the 9000 series 300.
• Michael Hayes for various thankless work he’s done trying to get the c30/c40 ports functional. Lots of loop and unroll improvements and fixes.
• Dara Hazeghi for wading through myriads of target-specific bug reports.
• Kate Hedstrom for staking the G77 folks with an initial testsuite.
• Richard Henderson for his ongoing SPARC, alpha, ia32, and ia64 work, loop opts, and generally fixing lots of old problems we’ve ignored for years, flow rewrite and lots of further stuff, including reviewing tons of patches.
• Aldy Hernandez for working on the PowerPC port, SIMD support, and various fixes.
• Nobuyuki Hikichi of Software Research Associates, Tokyo, contributed the support for the Sony NEWS machine.
• Kazu Hirata for caring and feeding the Renesas H8/300 port and various fixes.
• Katherine Holcomb for work on GNU Fortran.
• Manfred Hollstein for his ongoing work to keep the m88k alive, lots of testing and bug fixing, particularly of GCC configury code.
• Steve Holmgren for MachTen patches.
• Jan Hubicka for his x86 port improvements.
• Falk Hueffner for working on C and optimization bug reports.
• Bernardo Innocenti for his m68k work, including merging of ColdFire improvements and uClinux support.
• Christian Iseli for various bug fixes.
• Kamil Iskra for general m68k hacking.
• Lee Iverson for random fixes and MIPS testing.
• Andreas Jaeger for testing and benchmarking of GCC and various bug fixes.
Contributors to GCC

- Jakub Jelinek for his SPARC work and sibling call optimizations as well as lots of bug fixes and test cases, and for improving the Java build system.
- Janis Johnson for ia64 testing and fixes, her quality improvement sidetracks, and web page maintenance.
- Kean Johnston for SCO OpenServer support and various fixes.
- Tim Josling for the sample language treelang based originally on Richard Kenner’s “toy” language.
- Nicolai Josuttis for additional libstdc++ documentation.
- Klaus Kaempf for his ongoing work to make alpha-vms a viable target.
- Steven G. Kargl for work on GNU Fortran.
- David Kashtan of SRI adapted GCC to VMS.
- Ryszard Kabatek for many, many libstdc++ bug fixes and optimizations of strings, especially member functions, and for auto_ptr fixes.
- Geoffrey Keating for his ongoing work to make the PPC work for GNU/Linux and his automatic regression tester.
- Brendan Kehoe for his ongoing work with G++ and for a lot of early work in just about every part of libstdc++.
- Oliver M. Kellogg of Deutsche Aerospace contributed the port to the MIL-STD-1750A.
- Richard Kenner of the New York University Ultracomputer Research Laboratory wrote the machine descriptions for the AMD 29000, the DEC Alpha, the IBM RT PC, and the IBM RS/6000 as well as the support for instruction attributes. He also made changes to better support RISC processors including changes to common subexpression elimination, strength reduction, function calling sequence handling, and condition code support, in addition to generalizing the code for frame pointer elimination and delay slot scheduling. Richard Kenner was also the head maintainer of GCC for several years.
- Mumit Khan for various contributions to the Cygwin and Mingw32 ports and maintaining binary releases for Microsoft Windows hosts, and for massive libstdc++ porting work to Cygwin/Mingw32.
- Robin Kirkham for cpu32 support.
- Mark Klein for PA improvements.
- Thomas Koenig for various bug fixes.
- Bruce Korb for the new and improved fixincludes code.
- Benjamin Kosnik for his G++ work and for leading the libstdc++-v3 effort.
- Charles LaBrec contributed the support for the Integrated Solutions 68020 system.
- Asher Langton and Mike Kumbera for contributing Cray pointer support to GNU Fortran, and for other GNU Fortran improvements.
- Jeff Law for his direction via the steering committee, coordinating the entire egcs project and GCC 2.95, rolling out snapshots and releases, handling merges from GCC2, reviewing tons of patches that might have fallen through the cracks else, and random but extensive hacking.
- Marc Lehmann for his direction via the steering committee and helping with analysis and improvements of x86 performance.
• Victor Leikehman for work on GNU Fortran.
• Ted Lemon wrote parts of the RTL reader and printer.
• Kriang Lerdswanakij for C++ improvements including template as template parameter support, and many C++ fixes.
• Warren Levy for tremendous work on libgcj (Java Runtime Library) and random work on the Java front end.
• Alain Lichnewsky ported GCC to the MIPS CPU.
• Oskar Liljeblad for hacking on AWT and his many Java bug reports and patches.
• Robert Lipe for OpenServer support, new testsuites, testing, etc.
• Chen Liqin for various S+core related fixes/improvement, and for maintaining the S+core port.
• Weiwen Liu for testing and various bug fixes.
• Manuel López-Ibáñez for improving ‘-Wconversion’ and many other diagnostics fixes and improvements.
• Dave Love for his ongoing work with the Fortran front end and runtime libraries.
• Martin von Löwis for internal consistency checking infrastructure, various C++ improvements including namespace support, and tons of assistance with libstdc++/compiler merges.
• H.J. Lu for his previous contributions to the steering committee, many x86 bug reports, prototype patches, and keeping the GNU/Linux ports working.
• Greg McGary for random fixes and (someday) bounded pointers.
• Andrew MacLeod for his ongoing work in building a real EH system, various code generation improvements, work on the global optimizer, etc.
• Vladimir Makarov for hacking some ugly i960 problems, PowerPC hacking improvements to compile-time performance, overall knowledge and direction in the area of instruction scheduling, and design and implementation of the automaton based instruction scheduler.
• Bob Manson for his behind the scenes work on dejagnu.
• Philip Martin for lots of libstdc++ string and vector iterator fixes and improvements, and string clean up and testsuites.
• All of the Mauve project contributors, for Java test code.
• Bryce McKinlay for numerous GCJ and libgcj fixes and improvements.
• Adam Megacz for his work on the Microsoft Windows port of GCJ.
• Michael Meissner for LRS framework, ia32, m32r, v850, m88k, MIPS, powerpc, haifa, ECOFF debug support, and other assorted hacking.
• Jason Merrill for his direction via the steering committee and leading the G++ effort.
• Martin Michlmayr for testing GCC on several architectures using the entire Debian archive.
• David Miller for his direction via the steering committee, lots of SPARC work, improvements in jump.c and interfacing with the Linux kernel developers.
• Gary Miller ported GCC to Charles River Data Systems machines.
• Alfred Minarik for libstdc++ string and ios bug fixes, and turning the entire libstdc++
testsuite namespace-compatible.
• Mark Mitchell for his direction via the steering committee, mountains of C++ work,
load/store hoisting out of loops, alias analysis improvements, ISO C restrict support,
and serving as release manager for GCC 3.x.
• Alan Modra for various GNU/Linux bits and testing.
• Toon Moene for his direction via the steering committee, Fortran maintenance, and his
ongoing work to make us make Fortran run fast.
• Jason Molenda for major help in the care and feeding of all the services on the
gcc.gnu.org (formerly egcs.cygnus.com) machine—mail, web services, ftp services, etc
etc. Doing all this work on scrap paper and the backs of envelopes would have been . . .
difficult.
• Catherine Moore for fixing various ugly problems we have sent her way, including the
haifa bug which was killing the Alpha & PowerPC Linux kernels.
• Mike Moreton for his various Java patches.
• David Mosberger-Tang for various Alpha improvements, and for the initial IA-64 port.
• Stephen Moshier contributed the floating point emulator that assists in cross-
compilation and permits support for floating point numbers wider than 64 bits and
for ISO C99 support.
• Bill Moyer for his behind the scenes work on various issues.
• Philippe De Muyter for his work on the m68k port.
• Joseph S. Myers for his work on the PDP-11 port, format checking and ISO C99
support, and continuous emphasis on (and contributions to) documentation.
• Nathan Myers for his work on libstdc++-v3: architecture and authorship through the
first three snapshots, including implementation of locale infrastructure, string, shadow
C headers, and the initial project documentation (DESIGN, CHECKLIST, and so
forth). Later, more work on MT-safe string and shadow headers.
• Felix Natter for documentation on porting libstdc++.
• Nathanael Nerode for cleaning up the configuration/build process.
• NeXT, Inc. donated the front end that supports the Objective-C language.
• Hans-Peter Nilsson for the CRIS and MMIX ports, improvements to the search engine
setup, various documentation fixes and other small fixes.
• Geoff Noer for his work on getting cygwin native builds working.
• Diego Novillo for his work on Tree SSA, OpenMP, SPEC performance tracking web
pages, GIMPLE tuples, and assorted fixes.
• David O’Brien for the FreeBSD/alpha, FreeBSD/AMD x86-64, FreeBSD/ARM,
FreeBSD/PowerPC, and FreeBSD/SPARC64 ports and related infrastructure
improvements.
• Alexandre Oliva for various build infrastructure improvements, scripts and amazing
testing work, including keeping libtool issues sane and happy.
• Stefan Olsson for work on mt_alloc.
• Melissa O’Neill for various NeXT fixes.
• Rainer Orth for random MIPS work, including improvements to GCC’s o32 ABI support, improvements to dejagnu’s MIPS support, Java configuration clean-ups and porting work, and maintaining the IRIX, Solaris 2, and Tru64 UNIX ports.
• Hartmut Penner for work on the s390 port.
• Paul Petersen wrote the machine description for the Alliant FX/8.
• Alexandre Petit-Bianco for implementing much of the Java compiler and continued Java maintainership.
• Matthias Pfaller for major improvements to the NS32k port.
• Gerald Pfeifer for his direction via the steering committee, pointing out lots of problems we need to solve, maintenance of the web pages, and taking care of documentation maintenance in general.
• Andrew Pinski for processing bug reports by the dozen.
• Ovidiu Predescu for his work on the Objective-C front end and runtime libraries.
• Jerry Quinn for major performance improvements in C++ formatted I/O.
• Ken Raeburn for various improvements to checker, MIPS ports and various cleanups in the compiler.
• Rolf W. Rasmussen for hacking on AWT.
• David Reese of Sun Microsystems contributed to the Solaris on PowerPC port.
• Volker Reichelt for keeping up with the problem reports.
• Joern Rennecke for maintaining the sh port, loop, regmove & reload hacking.
• Loren J. Rittle for improvements to libstdc++-v3 including the FreeBSD port, threading fixes, thread-related configury changes, critical threading documentation, and solutions to really tricky I/O problems, as well as keeping GCC properly working on FreeBSD and continuous testing.
• Craig Rodrigues for processing tons of bug reports.
• Ola Rönnerup for work on mt_alloc.
• Gavin Romig-Koch for lots of behind the scenes MIPS work.
• David Ronis inspired and encouraged Craig to rewrite the G77 documentation in texinfo format by contributing a first pass at a translation of the old ‘g77-0.5.16/f/D0C’ file.
• Ken Rose for fixes to GCC’s delay slot filling code.
• Paul Rubin wrote most of the preprocessor.
• Pétur Runólfssson for major performance improvements in C++ formatted I/O and large file support in C++ filebuf.
• Chip Salzenberg for libstdc++ patches and improvements to locales, traits, Makefiles, libio, libtool hackery, and “long long” support.
• Juha Sarlin for improvements to the H8 code generator.
• Greg Satz assisted in making GCC work on HP-UX for the 9000 series 300.
• Roger Sayle for improvements to constant folding and GCC’s RTL optimizers as well as for fixing numerous bugs.
• Bradley Schatz for his work on the GCJ FAQ.
• Peter Schauer wrote the code to allow debugging to work on the Alpha.
• William Schelter did most of the work on the Intel 80386 support.
• Tobias Schlüter for work on GNU Fortran.
• Bernd Schmidt for various code generation improvements and major work in the reload pass as well a serving as release manager for GCC 2.95.3.
• Peter Schmid for constant testing of libstdc++—especially application testing, going above and beyond what was requested for the release criteria—and libstdc++ header file tweaks.
• Jason Schroeder for jcf-dump patches.
• Andreas Schwab for his work on the m68k port.
• Lars Segerlund for work on GNU Fortran.
• Dodji Seketeli for numerous C++ bug fixes and debug info improvements.
• Joel Sherrill for his direction via the steering committee, RTEMS contributions and RTEMS testing.
• Nathan Sidwell for many C++ fixes/improvements.
• Jeffrey Siegal for helping RMS with the original design of GCC, some code which handles the parse tree and RTL data structures, constant folding and help with the original VAX & m68k ports.
• Kenny Simpson for prompting libstdc++ fixes due to defect reports from the LWG (thereby keeping GCC in line with updates from the ISO).
• Franz Sirl for his ongoing work with making the PPC port stable for GNU/Linux.
• Andrei Slepuhin for assorted AIX hacking.
• Trevor Smigiel for contributing the SPU port.
• Christopher Smith did the port for Convex machines.
• Danny Smith for his major efforts on the Mingw (and Cygwin) ports.
• Randy Smith finished the Sun FPA support.
• Scott Snyder for queue, iterator, istream, and string fixes and libstdc++ testsuite entries. Also for providing the patch to G77 to add rudimentary support for INTEGER*1, INTEGER*2, and LOGICAL*1.
• Brad Spencer for contributions to the GLIBCPP_FORCE_NEW technique.
• Richard Stallman, for writing the original GCC and launching the GNU project.
• Jan Stein of the Chalmers Computer Society provided support for Genix, as well as part of the 32000 machine description.
• Nigel Stephens for various mips16 related fixes/improvements.
• Jonathan Stone wrote the machine description for the Pyramid computer.
• Graham Stott for various infrastructure improvements.
• John Stracke for his Java HTTP protocol fixes.
• Mike Stump for his Elxsi port, G++ contributions over the years and more recently his vxworks contributions
• Jeff Sturm for Java porting help, bug fixes, and encouragement.
• Shigeya Suzuki for this fixes for the bsd4 platforms.
- Ian Lance Taylor for the Go frontend, the initial mips16 and mips64 support, general configure hacking, fixincludes, etc.
- Holger Teutsch provided the support for the Clipper CPU.
- Gary Thomas for his ongoing work to make the PPC work for GNU/Linux.
- Philipp Thomas for random bug fixes throughout the compiler
- Jason Thorpe for thread support in libstdc++ on NetBSD.
- Kresten Krab Thorup wrote the run time support for the Objective-C language and the fantastic Java bytecode interpreter.
- Michael Tiemann for random bug fixes, the first instruction scheduler, initial C++ support, function integration, NS32k, SPARC and M88k machine description work, delay slot scheduling.
- Andreas Tobler for his work porting libgcj to Darwin.
- Teemu Torma for thread safe exception handling support.
- Leonard Tower wrote parts of the parser, RTL generator, and RTL definitions, and of the VAX machine description.
- Daniel Towner and Hariharan Sandanagobalane contributed and maintain the picoChip port.
- Tom Tromey for internationalization support and for his many Java contributions and libgcj maintainership.
- Lassi Tuura for improvements to config.guess to determine HP processor types.
- Petter Urkedal for libstdc++ CXXFLAGS, math, and algorithms fixes.
- Andy Vaught for the design and initial implementation of the GNU Fortran front end.
- Brent Verner for work with the libstdc++ cshadow files and their associated configure steps.
- Todd Vierling for contributions for NetBSD ports.
- Jonathan Wakely for contributing libstdc++ Doxygen notes and XHTML guidance.
- Dean Wakerley for converting the install documentation from HTML to texinfo in time for GCC 3.0.
- Krister Walfridsson for random bug fixes.
- Feng Wang for contributions to GNU Fortran.
- Stephen M. Webb for time and effort on making libstdc++ shadow files work with the tricky Solaris 8+ headers, and for pushing the build-time header tree.
- John Wehle for various improvements for the x86 code generator, related infrastructure improvements to help x86 code generation, value range propagation and other work, WE32k port.
- Ulrich Weigand for work on the s390 port.
- Zack Weinberg for major work on cpplib and various other bug fixes.
- Matt Welsh for help with Linux Threads support in GCJ.
- Urban Widmark for help fixing java.io.
- Mark Wielaard for new Java library code and his work integrating with Classpath.
- Dale Wiles helped port GCC to the Tahoe.
• Bob Wilson from Tensilica, Inc. for the Xtensa port.
• Jim Wilson for his direction via the steering committee, tackling hard problems in various places that nobody else wanted to work on, strength reduction and other loop optimizations.
• Paul Woegerer and Tal Agmon for the CRX port.
• Carlo Wood for various fixes.
• Tom Wood for work on the m88k port.
• Canqun Yang for work on GNU Fortran.
• Masanobu Yuhara of Fujitsu Laboratories implemented the machine description for the Tron architecture (specifically, the Gmicro).
• Kevin Zachmann helped port GCC to the Tahoe.
• Ayal Zaks for Swing Modulo Scheduling (SMS).
• Xiaoqiang Zhang for work on GNU Fortran.
• Gilles Zunino for help porting Java to Irix.

The following people are recognized for their contributions to GNAT, the Ada front end of GCC:
• Bernard Banner
• Romain Berrendonner
• Geert Bosch
• Emmanuel Briot
• Joel Brobecker
• Ben Brosgol
• Vincent Celier
• Arnaud Charlet
• Chien Chieng
• Cyrille Comar
• Cyrille Crozes
• Robert Dewar
• Gary Dismukes
• Robert Duff
• Ed Falis
• Ramon Fernandez
• Sam Figueroa
• Vasily Fofanov
• Michael Friess
• Franco Gasperoni
• Ted Giering
• Matthew Gingell
• Laurent Guerby
The following people are recognized for their contributions of new features, bug reports, testing and integration of classpath/libgcj for GCC version 4.1:

- Lillian Angel for JTree implementation and lots Free Swing additions and bug fixes.
- Wolfgang Baer for GapContent bug fixes.
- Anthony Balkissoon for JList, Free Swing 1.5 updates and mouse event fixes, lots of Free Swing work including JTable editing.
- Stuart Ballard for RMI constant fixes.
- Goffredo Baroncelli for HTTPURLConnection fixes.
- Gary Benson for MessageFormat fixes.
- Daniel Bonniot for Serialization fixes.
- Chris Burdess for lots of gnu.xml and http protocol fixes, StAX and DOM xml:id support.
- Ka-Hing Cheung for TreePath and TreeSelection fixes.
- Archie Cobbs for build fixes, VM interface updates, URLClassLoader updates.
- Kelley Cook for build fixes.
- Martin Cordova for Suggestions for better SocketTimeoutException.
Contributors to GCC

- David Daney for BitSet bug fixes, HttpURLConnection rewrite and improvements.
- Thomas Fitzsimmons for lots of upgrades to the gtk+ AWT and Cairo 2D support. Lots of imageio framework additions, lots of AWT and Free Swing bug fixes.
- Jeroen Frijters for ClassLoader and nio cleanups, serialization fixes, better Proxy support, bug fixes and IKVM integration.
- Santiago Gala for AccessControlContext fixes.
- Nicolas Geoffray for VMClassLoader and AccessController improvements.
- David Gilbert for basic and metal icon and plaf support and lots of documenting. Lots of Free Swing and metal theme additions. MetalIconFactory implementation.
- Anthony Green for MIDI framework, ALSA and DSSI providers.
- Andrew Haley for Serialization and URLClassLoader fixes, gcj build speedups.
- Kim Ho for JFileChooser implementation.
- Andrew John Hughes for Locale and net fixes, URI RFC2986 updates, Serialization fixes, Properties XML support and generic branch work, VMIntegration guide update.
- Bastiaan Huisman for TimeZone bug fixing.
- Andreas Jaeger for mprec updates.
- Paul Jenner for better `-Werror` support.
- Ito Kazumitsu for NetworkInterface implementation and updates.
- Roman Kennke for BoxLayout, GrayFilter and SplitPane, plus bug fixes all over. Lots of Free Swing work including styled text.
- Simon Kitching for String cleanups and optimization suggestions.
- Michael Koch for configuration fixes, Locale updates, bug and build fixes.
- Guilhem Lavaux for configuration, thread and channel fixes and Kaffe integration. JCL native Pointer updates. Logger bug fixes.
- David Lichteblau for JCL support library global/local reference cleanups.
- Aaron Luchko for JDWP updates and documentation fixes.
- Ziga Mahkovec for Graphics2D upgraded to Cairo 0.5 and new regex features.
- Sven de Marothy for BMP imageio support, CSS and TextLayout fixes. GtkImage rewrite, 2D, awt, free swing and date/time fixes and implementing the Qt4 peers.
- Casey Marshall for crypto algorithm fixes, FileChannel lock, SystemLogger and FileHandler rotate implementations, NIO FileChannel.map support, security and policy updates.
- Bryce McKinlay for RMI work.
- Audrius Meskauskas for lots of Free Corba, RMI and HTML work plus testing and documenting.
- Kalle Olavi Niemitalo for build fixes.
- Rainer Orth for build fixes.
- Andrew Overholt for File locking fixes.
- Ingo Proetel for Image, Logger and URLClassLoader updates.
- Olga Rodimina for MenuSelectionManager implementation.
Jan Roehrich for **BasicTreeUI** and **JTree** fixes.

Julian Scheid for documentation updates and gjdoc support.

Christian Schlichterle for zip fixes and cleanups.

Robert Schuster for documentation updates and beans fixes, **TreeNode** enumerations and **ActionCommand** and various fixes, XML and URL, AWT and Free Swing bug fixes.

Keith Seitz for lots of JDWP work.

Christian Thalinger for 64-bit cleanups, Configuration and VM interface fixes and **CACAO** integration, **fdlibm** updates.

Gael Thomas for **VMClassLoader** boot packages support suggestions.

Andreas Tobler for Darwin and Solaris testing and fixing, **Qt4** support for Darwin/OS X, **Graphics2D** support, **gtk+** updates.

Dalibor Topic for better **DEBUG** support, build cleanups and Kaffe integration. **Qt4** build infrastructure, **SHA1PRNG** and **GdkPixbugDecoder** updates.

Tom Tromey for Eclipse integration, generics work, lots of bug fixes and gcj integration including coordinating The Big Merge.

Mark Wielaard for bug fixes, packaging and release management, **Clipboard** implementation, system call interrupts and network timeouts and **GdkPixbufDecoder** fixes.

In addition to the above, all of which also contributed time and energy in testing GCC, we would like to thank the following for their contributions to testing:

- Michael Abd-El-Malek
- Thomas Arend
- Bonzo Armstrong
- Steven Ashe
- Chris Baldwin
- David Billinghurst
- Jim Blandy
- Stephane Bortzmeyer
- Horst von Brand
- Frank Braun
- Rodney Brown
- Sidney Cadot
- Bradford Castalia
- Robert Clark
- Jonathan Corbet
- Ralph Doncaster
- Richard Emberson
- Levente Farkas
- Graham Fawcett
- Mark Fernyhough
- Robert A. French
Contributors to GCC

- Jörgen Freyh
- Mark K. Gardner
- Charles-Antoine Gauthier
- Yung Shing Gene
- David Gilbert
- Simon Gornall
- Fred Gray
- John Griffin
- Patrik Hagglund
- Phil Hargett
- Amancio Hasty
- Takafumi Hayashi
- Bryan W. Headley
- Kevin B. Hendricks
- Joep Jansen
- Christian Joensson
- Michel Kern
- David Kidd
- Tobias Kuipers
- Anand Krishnaswamy
- A. O. V. Le Blanc
- llewelly
- Damon Love
- Brad Lucier
- Matthias Klose
- Martin Knoblauch
- Rick Lutowski
- Jesse Macnish
- Stefan Morrell
- Anon A. Mous
- Matthias Mueller
- Pekka Nikander
- Rick Niles
- Jon Olson
- Magnus Persson
- Chris Pollard
- Richard Polton
- Derk Reefman
- David Rees
- Paul Reilly
- Tom Reilly
- Torsten Rueger
- Danny Sadinoff
- Marc Schifer
- Erik Schnetter
- Wayne K. Schroll
- David Schuler
- Vin Shelton
- Tim Souder
- Adam Sulmicki
- Bill Thorson
- George Talbot
- Pedro A. M. Vazquez
- Gregory Warnes
- Ian Watson
- David E. Young
- And many others

And finally we’d like to thank everyone who uses the compiler, provides feedback and generally reminds us why we’re doing this work in the first place.
Option Index

GCC's command line options are indexed here without any initial ‘-’ or ‘--’. Where an option has both positive and negative forms (such as ‘-foption’ and ‘-fno-option’), relevant entries in the manual are indexed under the most appropriate form; it may sometimes be useful to look up both forms.

F
filtrans.............................. 594
filtrans-output-list............... 594
fwpa................................. 594

M
msoft-float.......................... 12
Concept Index

!  
‘!’ in constraint ................................. 307

#  
‘#’ in constraint ................................. 308
# in template .................................. 295
#pragma ......................................... 548

%  
‘%’ in constraint ................................. 308
% in GTY option ................................. 571
‘%’ in template ................................. 295

&  
‘&’ in constraint ................................. 307

(  
(nil) ........................................... 126

*  
‘*’ in constraint ................................. 308
* in template .................................. 296

+  
‘+’ in constraint ................................. 307

-  
‘-fsection-anchors’ ............................ 131, 475

/  
‘/’ in RTL dump ................................ 135
‘/f’ in RTL dump ................................ 135
‘/i’ in RTL dump ................................ 136
‘/j’ in RTL dump ................................ 136
‘/n’ in RTL dump ................................ 136
‘/o’ in RTL dump ................................ 136
‘/v’ in RTL dump ................................ 137

<  
‘<’ in constraint ................................. 302

=  
‘=’ in constraint ................................. 307

>  
‘>’ in constraint ................................. 303

?  
‘?’ in constraint ................................. 307

--ashrdi3 ........................................ 9
GNU Compiler Collection (GCC) Internals
650 GNU Compiler Collection (GCC) Internals
<table>
<thead>
<tr>
<th>Concept</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>fracms2 instruction pattern</td>
<td>345</td>
</tr>
<tr>
<td>fracunms2 instruction pattern</td>
<td>345</td>
</tr>
<tr>
<td>frame layout</td>
<td>436</td>
</tr>
<tr>
<td>FRAME_ADDR_RTX</td>
<td>437</td>
</tr>
<tr>
<td>FRAME_GROWS_DOWNWARD</td>
<td>436</td>
</tr>
<tr>
<td>FRAME_GROWS_DOWNWARD and virtual registers</td>
<td>145</td>
</tr>
<tr>
<td>FRAME_POINTER_CFA_OFFSET</td>
<td>439</td>
</tr>
<tr>
<td>frame_pointer_needed</td>
<td>459</td>
</tr>
<tr>
<td>FRAME_POINTER_REGNUM</td>
<td>443</td>
</tr>
<tr>
<td>FRAME_POINTER_REGNUM and virtual registers</td>
<td>145</td>
</tr>
<tr>
<td>frame_pointer_rtx</td>
<td>445</td>
</tr>
<tr>
<td>frame_related</td>
<td>135</td>
</tr>
<tr>
<td>function</td>
<td>206, 214</td>
</tr>
<tr>
<td>function call conventions</td>
<td>7</td>
</tr>
<tr>
<td>function entry and exit</td>
<td>458</td>
</tr>
<tr>
<td>function entry point, alternate function entry point</td>
<td>283</td>
</tr>
<tr>
<td>function properties</td>
<td>207</td>
</tr>
<tr>
<td>function-call insns</td>
<td>175</td>
</tr>
<tr>
<td>FUNCTION_ARG</td>
<td>449</td>
</tr>
<tr>
<td>FUNCTION_ARG_ADVANCE</td>
<td>452</td>
</tr>
<tr>
<td>FUNCTION_ARG_OFFSET</td>
<td>452</td>
</tr>
<tr>
<td>FUNCTION_ARG_PADDING</td>
<td>452</td>
</tr>
<tr>
<td>FUNCTION_ARG_REGNO_P</td>
<td>453</td>
</tr>
<tr>
<td>FUNCTION_BOUNDARY</td>
<td>406</td>
</tr>
<tr>
<td>FUNCTION_DECL</td>
<td>206, 214</td>
</tr>
<tr>
<td>FUNCTION_INCOMING_ARG</td>
<td>450</td>
</tr>
<tr>
<td>FUNCTION_MODE</td>
<td>547</td>
</tr>
<tr>
<td>FUNCTION_PROFILER</td>
<td>462</td>
</tr>
<tr>
<td>FUNCTION_TYPE</td>
<td>179</td>
</tr>
<tr>
<td>FUNCTION_VALUE</td>
<td>455</td>
</tr>
<tr>
<td>FUNCTION_VALUE_REGNO_P</td>
<td>456</td>
</tr>
<tr>
<td>functions, leaf</td>
<td>423</td>
</tr>
<tr>
<td>fundamental type</td>
<td>179</td>
</tr>
</tbody>
</table>

G

<table>
<thead>
<tr>
<th>Concept</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>'g' in constraint</td>
<td>303</td>
</tr>
<tr>
<td>'G' in constraint</td>
<td>303</td>
</tr>
<tr>
<td>garbage collector, invocation</td>
<td>577</td>
</tr>
<tr>
<td>garbage collector, troubleshooting</td>
<td>577</td>
</tr>
<tr>
<td>GCC and portability</td>
<td>5</td>
</tr>
<tr>
<td>GCC_DRIVER_HOST_INITIALIZATION</td>
<td>561</td>
</tr>
<tr>
<td>gcover_type</td>
<td>284</td>
</tr>
<tr>
<td>ge</td>
<td>155</td>
</tr>
<tr>
<td>ge and attributes</td>
<td>374</td>
</tr>
<tr>
<td>GE_EXPR</td>
<td>192</td>
</tr>
<tr>
<td>GEN_ERRNO_RTX</td>
<td>469</td>
</tr>
<tr>
<td>gencodes</td>
<td>120</td>
</tr>
<tr>
<td>general_operand</td>
<td>299</td>
</tr>
<tr>
<td>GENERAL_REGS</td>
<td>425</td>
</tr>
</tbody>
</table>

Generated files                      | 576      |

Generating assembler output          | 285      |

Generating insns                     | 291      |

GENERIC                                | 111, 177 |

generic predicates                     | 208      |

genflags                               | 120      |

get_attr                                | 374      |

get_attr_length                        | 378      |

GET_CLASS_NARROWEST_MODE               | 142      |

GET_CODE                                | 125      |

get_frame_code                         | 446      |

get_insns                               | 166      |

get_last_insn                          | 166      |

GET_MODE                                | 141      |

GET_MODE_ALIGNMENT                      | 142      |

GET_MODE_BITSIZE                        | 141      |

GET_MODE_CLASS                          | 141      |

GET_MODE_FBIT                           | 142      |

GET_MODE_IBIT                           | 142      |

GET_MODE_MASK                           | 142      |

GET_MODE_NAME                           | 141      |

GET_MODE_NUMITS                         | 142      |

GET_MODE_SIZE                           | 141      |

GET_MODE_UNIT_SIZE                      | 142      |

GET_MODE_WIDER_MODE                     | 141      |

GET_RTX_CLASS                           | 126      |

GET_RTX_FORMAT                          | 128      |

GET_RTX_LENGTH                          | 128      |

geu                                      | 155      |

geu and attributes                      | 374      |

GCC                                      | 571      |

gcc_collect                             | 577      |

CIMPLE                                   | 111, 112, 221 |
CIMPLE_Exception Handling                | 225      |
CIMPLE instruction set                   | 225      |
CIMPLE sequences                         | 250      |

gimple_addresses_taken                  | 231      |

CINPLE_ASM                               | 233      |

gimple_asm_clear_volatile               | 234      |

gimple_asm_clobber_ops                  | 233      |

gimple_asm_input_ops                    | 233      |

gimple_asm_nclobbers                    | 233      |

gimple_asm_ninputs                      | 233      |

gimple_asm_noutputs                     | 233      |

gimple_asm_output_ops                   | 233      |

gimple_asm_set_clobber_ops              | 233      |

gimple_asm_set_input_ops                | 233      |

gimple_asm_set_output_ops               | 233      |

gimple_asm_set_volatile                 | 234      |

gimple_asm_string                       | 233      |

gimple_asm_volatile_p                   | 233      |

CINPLE_ASSIGN                            | 234      |

gimple_assign_cast_p                    | 230, 235|

gimple_assign_lhs                       | 234      |

gimple_assign_lhs_ptr                   | 234      |

gimple_assign_rhs_class                 | 234      |

gimple_assign_rhs_code                  | 234      |

gimple_assign_rsh                        | 234     |
<table>
<thead>
<tr>
<th>Concept Index</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>HAS_LONG_COND_BRANCH</td>
<td>542</td>
</tr>
<tr>
<td>HAS_LONG_UNCOND_BRANCH</td>
<td>542</td>
</tr>
<tr>
<td>HAVE_DOS_BASED_FILE_SYSTEM</td>
<td>560</td>
</tr>
<tr>
<td>HAVE_POST_DECREMENT</td>
<td>469</td>
</tr>
<tr>
<td>HAVE_POST_INCREMENT</td>
<td>469</td>
</tr>
<tr>
<td>HAVE_POST_MODIFY DISP</td>
<td>469</td>
</tr>
<tr>
<td>HAVE_POST_MODIFY_REG</td>
<td>469</td>
</tr>
<tr>
<td>HAVE_PRE_DECREMENT</td>
<td>469</td>
</tr>
<tr>
<td>HAVE_PRE_INCREMENT</td>
<td>469</td>
</tr>
<tr>
<td>HAVE_PRE_MODIFY DISP</td>
<td>469</td>
</tr>
<tr>
<td>HAVE_PRE_MODIFY_REG</td>
<td>469</td>
</tr>
<tr>
<td>HCmode</td>
<td>140</td>
</tr>
<tr>
<td>HFmode</td>
<td>138</td>
</tr>
<tr>
<td>high</td>
<td>144</td>
</tr>
<tr>
<td>HImode</td>
<td>138</td>
</tr>
<tr>
<td>HImode, in insn</td>
<td>170</td>
</tr>
<tr>
<td>HONOR_REG_ALLOC_ORDER</td>
<td>421</td>
</tr>
<tr>
<td>host configuration</td>
<td>559</td>
</tr>
<tr>
<td>host functions</td>
<td>559</td>
</tr>
<tr>
<td>host hooks</td>
<td>559</td>
</tr>
<tr>
<td>host makefile fragment</td>
<td>565</td>
</tr>
<tr>
<td>HOST_BIT_BUCKET</td>
<td>560</td>
</tr>
<tr>
<td>HOST_EXECUTABLE_SUFFIX</td>
<td>560</td>
</tr>
<tr>
<td>HOST_HOOKS_EXTRA_SIGNALS</td>
<td>559</td>
</tr>
<tr>
<td>HOST_HOOKS_GT_PCH_ALLOC_GRANULARITY</td>
<td>559</td>
</tr>
<tr>
<td>HOST_HOOKS_GT_PCH_GET_ADDRESS</td>
<td>559</td>
</tr>
<tr>
<td>HOST_HOOKS_GT_PCH_USE_ADDRESS</td>
<td>559</td>
</tr>
<tr>
<td>HOST_LACKS_INODE_NUMBERS</td>
<td>561</td>
</tr>
<tr>
<td>HOST_LONG_FORMAT</td>
<td>562</td>
</tr>
<tr>
<td>HOST_LONG_LONG_FORMAT</td>
<td>562</td>
</tr>
<tr>
<td>HOST_OBJECT_SUFFIX</td>
<td>560</td>
</tr>
<tr>
<td>HOST_PTR_PRINTF</td>
<td>562</td>
</tr>
<tr>
<td>HOST_TEXT_SECTION_NAME</td>
<td>492</td>
</tr>
<tr>
<td>HQmode</td>
<td>139</td>
</tr>
</tbody>
</table>

**I**

*i* in constraint | 303 |
**I** in constraint | 303 |
identifier | 179 |
IDENTIFIER_LENGTH | 179 |
IDENTIFIER_NODE | 179 |
IDENTIFIER_OPNAME_P | 179 |
IDENTIFIER_POINTER | 179 |
IDENTIFIER_TYPENAME_P | 179 |
IEEE 754-2008 | 16 |
IP_COND | 217 |
if_marked | 574 |
IP_STMT | 217 |
if_then else | 155 |
if_then else and attributes | 373 |
if_then else usage | 159 |
IFCVT_EXTRA_FIELDS | 551 |
IFCVT_INIT_EXTRA_FIELDS | 551 |
IFCVT_MODIFY_CANCEL | 551 |
IFCVT_MODIFY_FINAL | 551 |
IFCVT_MODIFY_insn | 551 |
IFCVT_MODIFY_MULTIPLE_TESTS | 550 |
IFCVT_MODIFY_TESTS | 550 |
IMAPART_EXPR | 192 |
Immediate Uses | 261 |
Immediate operand | 298 |
IMMEDIATE_PREFIX | 518 |
in_struct | 136 |
in_struct, in code_label and note | 132 |
in_struct, in insn and jump_insn and call_insn | 132 |
instruct, in insn, jump_insn and call_insn | 134 |
instruct, in mem | 132 |
instruct, in subreg | 135 |
include | 367 |
INCLUDED_DEFAULTS | 399 |
inclusive-or, bitwise | 153 |
INCOMING_FRAME_SP_OFFSET | 439 |
INCOMING_REGNO | 420 |
INCOMING_RETURN_ADDR_RTX | 438 |
INCOMING_STACK_BOUNDARY | 406 |
INDEX_REG_CLASS | 427 |
direct_jump instruction pattern | 349 |
direct_operand | 299 |
DIRECTREF | 191 |
INIT_ARRAY_SECTION_ASM_OP | 493 |
INIT_CUMULATIVE_ARGS | 451 |
INIT_CUMULATIVE_INCLUDING_ARGS | 452 |
INIT_CUMULATIVE_LIBCALL_ARGS | 451 |
INIT_ENVIRONMENT | 398 |
INIT_EXPANDERS | 404 |
INIT_EXP | 192 |
init_machine_status | 404 |
init_one_func | 468 |
INIT_SECTION_ASM_OP | 492, 514 |
INITIAL_ELIMINATION_ASM_OP | 447 |
INITIAL_FRAME_ADDRESS_RTX | 437 |
INITIAL_FRAME_POINTER | 446 |
initialization routines | 512 |
linenumbering | 536 |
insert_insn_on_edge | 286 |
insn | 167 |
insn and '/f' | 133 |
insn and '/j' | 134 |
insn and '/s' | 132, 134 |
insn and '/u' | 132 |
insn and '/v' | 132 |
insn attributes | 371 |
insn canonicalization | 360 |
insn includes | 367 |
insn lengths, computing | 378 |
insn splitting | 364 |
insn_attr.h | 372 |
INSN_ANNULLED_BRANCH | 132 |
INSN_CODE | 170 |
INSN_DELETED_P | 132 |
INSN_FROM_TARGET | 132 |
insn_list | 174 |
INSN_REFERENCES_ARE_DELAYED | 549 |
IV analysis on GIMPLE...

jump instruction pattern               347
jump instruction patterns             357
jump instructions and set           159
jump, in call_insn                  134
jump, in insn                        134
jump, in mem                        133
JUMP_ALIGN                          523
jump_insn                           167
jump_insn and '/f'                133
jump_insn and '/s'                192, 134
jump_insn and '/u'                132
jump_insn and '/v'                132
JUMP_LABEL                          167
JUMP_TABLES_IN_TEXT_SECTION       493
Jumps                                203

L

LABEL_ALIGN                       524
LABEL_ALIGN_After_BARRIER        523
LABEL_ALT_ENTRY_P               418
LABEL_ALTERNATE_NAME            283
LABEL_DECL                        184
LABEL_KIND                       168
LABEL_NUSES                      168
LABEL_Preserve_P                132
label_ref                         143
label_ref and '/f'              132
label_ref and '/s'              176
label_ref, RTL sharing           176
LABEL_REF_NONLOCAL_P            132
lang_hooks.gimplify_expr        112
lang_hooks.parse_file           111
glanguage-dependent trees        208
glanguage-independent intermediate representation  111

glarge return values              457
GLARGEST_EXPONENT_IS_NORMAL      411
GLAST_STACK_REG                  424
GLAST_VIRTUAL_REGISTER         145
1ceil2n2                         342
LCSSA                             272
LD_FINI_SWITCH                   514
LD_INIT_SWITCH                   514
LDD_SUFFIX                       516
le                                 155
le and attributes               374
LE_EXPR                            192
leaf functions                   423
leaf_function_p                 348
LEAF_REG_REMAP                   424
LEAF_REGISTERS                   424
left rotate                       153
left shift                        153
LEGITIMATE_CONSTANT_P           473
LEGITIMATE_PIC_OPERAND_P        496
LEGITIMIZE_RELOAD_ADDRESS     471
<table>
<thead>
<tr>
<th>Concept Index 665</th>
</tr>
</thead>
<tbody>
<tr>
<td>length .................. 572</td>
</tr>
<tr>
<td>less than .............. 155</td>
</tr>
<tr>
<td>less than or equal ...... 155</td>
</tr>
<tr>
<td>leu ........................ 155</td>
</tr>
<tr>
<td>leu and attributes ...... 374</td>
</tr>
<tr>
<td>lflooren2 .................. 342</td>
</tr>
<tr>
<td>L SPEC ............................ 395</td>
</tr>
<tr>
<td>LIB2FUNCS_EXTRA ............ 563</td>
</tr>
<tr>
<td>LIBCALL_VALUE .............. 455</td>
</tr>
<tr>
<td>'libgcc.a' ................. 468</td>
</tr>
<tr>
<td>LIBGCC_SPEC .............. 395</td>
</tr>
<tr>
<td>LIBGCC2_CFLAGS ........... 563</td>
</tr>
<tr>
<td>LIBGCC2_HAS_DF_MODE ........ 414</td>
</tr>
<tr>
<td>LIBGCC2_HAS_TP_MODE ...... 415</td>
</tr>
<tr>
<td>LIBGCC2_HAS_XF_MODE .... 415</td>
</tr>
<tr>
<td>LIBGCC2_LONG_DOUBLE_TYPE_SIZE .... 414</td>
</tr>
<tr>
<td>LIBGCC2_UNWIND_ATTRIBUTE ... 557</td>
</tr>
<tr>
<td>library subroutine names ... 468</td>
</tr>
<tr>
<td>LIBRARY_PATH_ENV ........... 550</td>
</tr>
<tr>
<td>LIMIT_RELOAD_CLASS ........ 429</td>
</tr>
<tr>
<td>Linear loop transformations framework 277</td>
</tr>
<tr>
<td>LINK_COMMAND_SPEC ......... 397</td>
</tr>
<tr>
<td>LINK_EH_SPEC ............. 396</td>
</tr>
<tr>
<td>LINK_ELIMINATE_DUPLICATE_LDIRECTORIES ...... 397</td>
</tr>
<tr>
<td>LINK_GCC_C_SEQUENCE_SPEC ...... 397</td>
</tr>
<tr>
<td>LINK_LIBGCC_SPECIAL_1 ..... 397</td>
</tr>
<tr>
<td>LINK_SPEC ............. 395</td>
</tr>
<tr>
<td>list .......................... 179</td>
</tr>
<tr>
<td>Liveness representation ... 286</td>
</tr>
<tr>
<td>lo_sum ..................... 150</td>
</tr>
<tr>
<td>load address instruction ... 304</td>
</tr>
<tr>
<td>LOAD_EXTEND_OP .......... 543</td>
</tr>
<tr>
<td>load_multiple instruction pattern 334</td>
</tr>
<tr>
<td>LOCAL_ALIGNMENT .......... 408</td>
</tr>
<tr>
<td>LOCAL_CLASS_P ............. 213</td>
</tr>
<tr>
<td>LOCAL_DECL_ALIGNMENT ....... 408</td>
</tr>
<tr>
<td>LOCAL_INCLUDE_DIR ......... 398</td>
</tr>
<tr>
<td>LOCAL_LABEL_PREFIX ........ 518</td>
</tr>
<tr>
<td>LOCAL_REGND .............. 420</td>
</tr>
<tr>
<td>LOG_LINKS .................. 171</td>
</tr>
<tr>
<td>Logical Operators .......... 227</td>
</tr>
<tr>
<td>logical-and, bitwise .... 153</td>
</tr>
<tr>
<td>log2 instruction pattern ... 340</td>
</tr>
<tr>
<td>LONG_ACCUM_TYPE_SIZE .... 414</td>
</tr>
<tr>
<td>LONG_DOUBLE_TYPE_SIZE .... 414</td>
</tr>
<tr>
<td>LONG_FRAC_TYPE_SIZE .... 414</td>
</tr>
<tr>
<td>LONG_LONG_ACCUM_TYPE_SIZE .... 414</td>
</tr>
<tr>
<td>LONG_LONG_FRAC_TYPE_SIZE .... 414</td>
</tr>
<tr>
<td>LONG_LONG_TYPE_SIZE .... 413</td>
</tr>
<tr>
<td>LONG_TYPE_SIZE ........... 413</td>
</tr>
<tr>
<td>longjmp and automatic variables ... 7</td>
</tr>
<tr>
<td>Loop analysis ............. 269</td>
</tr>
<tr>
<td>Loop manipulation .......... 272</td>
</tr>
<tr>
<td>Loop querying ............ 271</td>
</tr>
<tr>
<td>Loop representation ...... 269</td>
</tr>
<tr>
<td>Loop-closed SSA form ...... 272</td>
</tr>
<tr>
<td>LOOP_ALIGN ............. 523</td>
</tr>
<tr>
<td>LOOP_EXPR ............. 192</td>
</tr>
<tr>
<td>looping instruction patterns ..... 358</td>
</tr>
<tr>
<td>lowering, language-dependent intermediate representation .......... 111</td>
</tr>
<tr>
<td>lrintm2 .................... 341</td>
</tr>
<tr>
<td>lroundm2 .................... 341</td>
</tr>
<tr>
<td>LSHIFT_EXPR .................. 192</td>
</tr>
<tr>
<td>lshiftr .......................... 153</td>
</tr>
<tr>
<td>lshiftr and attributes ...... 374</td>
</tr>
<tr>
<td>lshr3 instruction pattern .. 339</td>
</tr>
<tr>
<td>lt .................................. 155</td>
</tr>
<tr>
<td>lt and attributes .......... 374</td>
</tr>
<tr>
<td>LT_EXPR ....................... 192</td>
</tr>
<tr>
<td>LTGT_EXPR ..................... 192</td>
</tr>
<tr>
<td>lto .................................. 587</td>
</tr>
<tr>
<td>ltrans .......................... 587</td>
</tr>
<tr>
<td>ltu .................................. 155</td>
</tr>
</tbody>
</table>

**M**

'm' in constraint ........................ 302
machine attributes .................... 534
machine description macros ........... 393
machine descriptions ................ 289
machine mode conversions .............. 157
machine modes ............. 137
machine specific constraints .......... 308
machine-independent predicates ...... 298
macros, target description ........... 393
maddm4 instruction pattern .......... 338
MAKE_DECL_ONE_ONLY ......... 508
make_phi_node .................. 247
make_safe_from ................. 363
makefile fragment ........... 563
makefile targets ................ 65
MALLOCA_ABI_ALIGNMENT .......... 406
Manipulating GIMPLE statements ...... 230
mark_hook ................ 574
marking roots ................ 576
MASK_RETURN_ADDR .......... 521
match_dup ...................... 292, 371
match_dup and attributes ........... 378
match_op_dup .................. 293
match_operand .................. 291
match_operand and attributes ...... 374
match_operator ................. 292
match_par_dup .................. 294
match_parallel ................. 294
match_scratch .................. 292, 371
matching constraint ............. 304
matching operands ............. 295
math library .......................... 12
math, in RTL .................. 150
MATH_LIBRARY ............. 550
matherr ................ 468
MAX_BITS_PER_WORD ............ 405
MAX_CONDITIONAL_EXECUTE ...... 550
MAX_FIXED_MODE_SIZE .......... 410
MAX_MOVE_MAX .................. 544
<table>
<thead>
<tr>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘n’ in constraint</td>
</tr>
<tr>
<td>N_REG_CLASSES</td>
</tr>
<tr>
<td>name</td>
</tr>
<tr>
<td>named address spaces</td>
</tr>
<tr>
<td>named patterns and conditions</td>
</tr>
<tr>
<td>names, pattern</td>
</tr>
<tr>
<td>namespace, scope</td>
</tr>
<tr>
<td>NAMESPACE_DECL</td>
</tr>
<tr>
<td>NATIVE_SYSTEM_HEADER_DIR</td>
</tr>
<tr>
<td>ne</td>
</tr>
<tr>
<td>ne and attributes</td>
</tr>
<tr>
<td>NE_EXPR</td>
</tr>
<tr>
<td>nearbyint2 instruction pattern</td>
</tr>
<tr>
<td>neg</td>
</tr>
<tr>
<td>neg and attributes</td>
</tr>
<tr>
<td>neg, canonicalization of instruction pattern</td>
</tr>
<tr>
<td>NEGATE_EXPR</td>
</tr>
<tr>
<td>negation</td>
</tr>
<tr>
<td>negation with signed saturation</td>
</tr>
<tr>
<td>negation with unsigned saturation</td>
</tr>
<tr>
<td>negm2 instruction pattern</td>
</tr>
<tr>
<td>nested functions, trampolines for 346</td>
</tr>
<tr>
<td>nested_ptr</td>
</tr>
<tr>
<td>next_bb, prev_bb, FOR_EACH_BB</td>
</tr>
<tr>
<td>NEXT_INSN</td>
</tr>
<tr>
<td>NEXT_OBJC_RUNTIME</td>
</tr>
<tr>
<td>nil</td>
</tr>
<tr>
<td>NM_FLAGS</td>
</tr>
<tr>
<td>NO_DBG_BNSYM_ENSYM</td>
</tr>
<tr>
<td>NO_DBG_FUNCTION_END</td>
</tr>
<tr>
<td>NO_DBG_GCC_MARKER</td>
</tr>
<tr>
<td>NO_DBG_MAIN_SOURCE_DIRECTORY</td>
</tr>
<tr>
<td>NO_DOLLAR_IN_LABEL</td>
</tr>
<tr>
<td>NO_DOT_IN_LABEL</td>
</tr>
<tr>
<td>NO_FUNCTION_CSE</td>
</tr>
<tr>
<td>NO_IMPLICIT_EXTERN_C</td>
</tr>
<tr>
<td>NO_PROFILE_COUNTERS</td>
</tr>
<tr>
<td>NO_REGS</td>
</tr>
<tr>
<td>NON_LVALUE_EXPR</td>
</tr>
<tr>
<td>nondeterministic finite state automaton</td>
</tr>
<tr>
<td>nonlabeled_operand</td>
</tr>
<tr>
<td>nonlocal goto handler</td>
</tr>
<tr>
<td>nonlocal goto instruction pattern</td>
</tr>
<tr>
<td>nonlocal goto receiver instruction pattern</td>
</tr>
<tr>
<td>nonmemery_operand</td>
</tr>
<tr>
<td>nonnullglobable memory reference</td>
</tr>
<tr>
<td>nop instruction pattern</td>
</tr>
<tr>
<td>NDP_EXPR</td>
</tr>
<tr>
<td>normal predicates</td>
</tr>
<tr>
<td>not</td>
</tr>
<tr>
<td>not and attributes</td>
</tr>
<tr>
<td>not equal</td>
</tr>
<tr>
<td>not, canonicalization of note</td>
</tr>
<tr>
<td>note</td>
</tr>
<tr>
<td>note and ‘/i’</td>
</tr>
<tr>
<td>note and ‘/v’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>O</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘o’ in constraint</td>
</tr>
<tr>
<td>OBJC_GEN_METHOD_LABEL</td>
</tr>
<tr>
<td>OBJC_JBLEN</td>
</tr>
<tr>
<td>OBJECT_FORMAT_COFF</td>
</tr>
<tr>
<td>OFFSET_TYPE</td>
</tr>
<tr>
<td>offsettable address</td>
</tr>
<tr>
<td>01mode</td>
</tr>
<tr>
<td>Omega a solver for linear programming problems</td>
</tr>
<tr>
<td>OMP_ATOMIC</td>
</tr>
<tr>
<td>OMP_CLAUSE</td>
</tr>
<tr>
<td>OMP_CONTINUE</td>
</tr>
<tr>
<td>OMP_CRITICAL</td>
</tr>
<tr>
<td>OMP_FOR</td>
</tr>
<tr>
<td>OMP_MASTER</td>
</tr>
<tr>
<td>OMP_ORDERED</td>
</tr>
<tr>
<td>OMP_PARALLEL</td>
</tr>
<tr>
<td>OMP_RETURN</td>
</tr>
<tr>
<td>OMP_SECTION</td>
</tr>
<tr>
<td>OMP_SECTIONS</td>
</tr>
<tr>
<td>OMP_SINGLE</td>
</tr>
<tr>
<td>one_cmplx2 instruction pattern</td>
</tr>
<tr>
<td>operand access</td>
</tr>
<tr>
<td>Operand Access Routines</td>
</tr>
<tr>
<td>operand constraints</td>
</tr>
<tr>
<td>Operand Iterators</td>
</tr>
<tr>
<td>operand predicates</td>
</tr>
<tr>
<td>operand substitution</td>
</tr>
<tr>
<td>operands</td>
</tr>
<tr>
<td>operands</td>
</tr>
<tr>
<td>Operands</td>
</tr>
<tr>
<td>operator predicates</td>
</tr>
<tr>
<td>‘optc-gen.awk’</td>
</tr>
<tr>
<td>Optimization infrastructure for GIMPLE</td>
</tr>
<tr>
<td>Symbol</td>
</tr>
<tr>
<td>------------------------------</td>
</tr>
<tr>
<td>OPTIMIZE_MODE_SWITCHING</td>
</tr>
<tr>
<td>option specification files</td>
</tr>
<tr>
<td>OPTION_DEFAULT SPECS</td>
</tr>
<tr>
<td>optional hardware or system features</td>
</tr>
<tr>
<td>options, directory search</td>
</tr>
<tr>
<td>order of register allocation</td>
</tr>
<tr>
<td>ordered_comparison_operator</td>
</tr>
<tr>
<td>ORDERED_EXPR</td>
</tr>
<tr>
<td>Ordering of Patterns</td>
</tr>
<tr>
<td>ORIGINAL_REGNO</td>
</tr>
<tr>
<td>other register constraints</td>
</tr>
<tr>
<td>OUTGOING_REG_PARM_STACK_SPACE</td>
</tr>
<tr>
<td>OUTGOING_REGNO</td>
</tr>
<tr>
<td>output of assembler code</td>
</tr>
<tr>
<td>output statements</td>
</tr>
<tr>
<td>output templates</td>
</tr>
<tr>
<td>OUTPUT_ADDR_CONST_EXTRA</td>
</tr>
<tr>
<td>output_asm_insn</td>
</tr>
<tr>
<td>OUTPUT_QUOTED_STRING</td>
</tr>
<tr>
<td>OVERLAPPING_REGISTER_NAMES</td>
</tr>
<tr>
<td>OVERLOAD</td>
</tr>
<tr>
<td>OVERRIDE_ABI_FORMAT</td>
</tr>
<tr>
<td>OVL_CURRENT</td>
</tr>
<tr>
<td>OVL_NEXT</td>
</tr>
</tbody>
</table>

**P**

- P`p` in constraint ........................................ 304
- PAD_VARARGS_DOWN ........................................... 452
- parallel ..................................................... 162
- param_is ..................................................... 573
- parameters, c++ abi ...................................... 539
- parameters, miscellaneous ................................ 542
- parameters, precompiled headers ......................... 538
- param_is ..................................................... 573
- parity ......................................................... 154
- parity2 instruction pattern ................................. 342
- PARM_BOUNDARY ............................................. 406
- PARM_DECL .................................................... 184
- PARSER_LDD_OUTPUT ......................................... 516
- passes and files of the compiler .......................... 111
- passing arguments ......................................... 7
- PATH_SEPARATOR .............................................. 560
- PATTERN ....................................................... 170
- pattern conditions ......................................... 290
- pattern names ............................................... 332
- Pattern Ordering ............................................ 357
- patterns ....................................................... 289
- pc .............................................................. 150
- pc and attributes ........................................... 378
- pc, RTL sharing ............................................. 176
- PC_REGNUM ..................................................... 420
- pc_rtx ......................................................... 150
- PCC_BITFIELD_TYPE_MATTERS ................................. 409
- PCC_STATIC_STRUCT_RETURN ................................. 457
- PDImode ....................................................... 138
- peephole optimization, RTL representation .............. 162
- peephole optimizer definitions ............................ 368
- per-function data .......................................... 403
- percent sign ................................................ 295
- PHI nodes ..................................................... 263
- PHI_ARG_DEF .................................................. 264
- PHI_ARG_EDGE ................................................ 264
- PHI_ARG_ELT .................................................. 263
- PHI_NUM_ARGS ................................................ 263
- PHI_RESULT ................................................... 263
- PIC ............................................................ 496
- PIC_OFFSET_TABLE_REG_CALL_CLOBBERED .................... 496
- PIC_OFFSET_TABLE_REGNUM .................................. 496
- pipeline hazard recognizer ................................. 380
- Plugins ....................................................... 579
- plus ........................................................... 150
- plus and attributes ........................................ 374
- plus, canonicalization of .................................. 360
- PLUS_EXPR ..................................................... 192
- Pmode ........................................................ 547
- pmode_register_operand ..................................... 298
- pointer ....................................................... 179
- POINTER_PLUS_EXPR .......................................... 192
- POINTER_SIZE ................................................ 405
- POINTER_TYPE ................................................. 179
- POINTERS_EXTEND_UNSIGNED .................................. 405
- pop_operand .................................................. 299
- popcount ...................................................... 154
- popcountm2 instruction pattern ............................ 342
- portability ................................................... 5
- position independent code ................................ 496
- post_dec ...................................................... 164
- post_inc ...................................................... 164
- post_modify .................................................. 164
- POSTDECREMENT_EXPR ........................................ 192
- POSTINCREMENT_EXPR ........................................ 192
- POWI_MAX_MULTS .............................................. 555
- powm3 instruction pattern .................................. 340
- pragma ......................................................... 548
- pre_dec ....................................................... 164
- pre_GCC3_DWARF_FRAME_REGISTERS ............................ 445
- pre_inc ....................................................... 164
- pre_modify ................................................... 164
- PREDECREMENT_EXPR .......................................... 192
- predefined macros .......................................... 400
- predicates .................................................... 297
- predicates and machine modes ............................... 297
- predication ................................................... 480
- predict.def .................................................. 283
- PREFERRED_DEBUGGING_TYPE ................................. 525
- PREFERRED_OUTPUT_RELOAD_CLASS ............................ 429
- PREFERRED_RELOAD_CLASS .................................... 428
- PREFERRED_STACK_BOUNDARY .................................. 406
- prefetch ..................................................... 164
- prefetch and '/v' ............................................. 135
- prefetch instruction pattern ............................... 354
- PREFETCH_SCHEDULE_BARRIER_P ................................ 135
- PREINCREMENT_EXPR .......................................... 192
- presence_set ................................................ 383
preserving SSA form ................................. 264
preserving virtual SSA form ......................... 265
prev_active_insn ................................ 369
PREV_INSN ............................................ 166
PRINT_OPERAND ................................. 517
PRINT_OPERAND_ADDRESS .......................... 518
PRINT_OPERAND_PUNCT_VALID_P .................. 518
probe_stack instruction pattern ............... 351
processor functional units ....................... 380, 381
processor pipeline description ................. 380
product ............................................. 152
profile feedback .................................. 283
profile representation .............................. 283
PROFILE_BEFORE_PROLOGUE ....................... 463
PROFILE_HOOK .................................... 462
profiling, code generation ......................... 462
program counter .................................. 150
prologue ............................................ 458
prologue instruction pattern ...................... 353
PROMOTE_MODE ..................................... 405
pseudo registers .................................. 144
PSI_mode ............................................ 138
PTRDIFF_TYPE ....................................... 416
purge_dead_edges ................................. 281, 286
push address instruction ......................... 304
PUSH_ARGS ......................................... 447
PUSH_ARGS_REVERSED .............................. 447
push_operand ....................................... 299
push_reload ........................................ 472
PUSH_ROUNDING ..................................... 447
pushm1 instruction pattern ....................... 335
PUT_CODE ............................................ 125
PUT_MODE ............................................ 141
PUT_REG_NOTE_KIND ................................ 171
PUT_SDB ............................................. 531

Q
QCmode .............................................. 140
QFmode .............................................. 138
QI_mode ............................................. 137
QI_mode, in insn .................................. 170
QQmode .............................................. 138
qualified type ..................................... 179, 209
querying function unit reservations .......... 381
question mark ..................................... 307
quotient ............................................. 152

R
‘r’ in constraint .................................. 303
RANGE_TEST_NON_SHORT_CIRCUIT ................. 484
REDIV_EXPR ......................................... 192
READONLY_DATA_SECTION_ASM_OP ................ 492
real operands ...................................... 257
REAL_ARITHMETIC .................................. 532
REAL_CST ............................................. 189
REAL_LIBGCC_SPEC ................................ 395
REAL_NM_FILE_NAME ................................ 515
REAL_TYPE ......................................... 179
REAL_VALUE_ABS .................................. 533
REAL_VALUE_ATOF ................................ 532
REAL_VALUE_FIX ................................... 532
REAL_VALUE_FROM_INT .............................. 533
REAL_VALUE_ISINF ................................ 532
REAL_VALUE_ISNAN ................................ 532
REAL_VALUE_NEGATE ................................ 533
REAL_VALUE_NEGATIVE .............................. 532
REAL_VALUE_TO_INT ................................ 533
REAL_VALUE_TO_TARGET_DECIMAL128 ............. 503
REAL_VALUE_TO_TARGET_DECIMAL32 ............... 503
REAL_VALUE_TO_TARGET_DECIMAL64 ............... 503
REAL_VALUE_TO_TARGET_DOUBLE ................... 502
REAL_VALUE_TO_TARGET_LONG_DOUBLE ............. 502
REAL_VALUE_TO_TARGET_SINGLE ................... 502
REAL_VALUE_TRUNCATE ................................ 533
REAL_VALUE_TYPE .................................. 532
REAL_VALUE_UNSIGNED_FIX ......................... 532
REAL_VALUES_EQUAL ................................ 532
REAL_VALUES_LESS ................................ 532
REALPART_EXPR ..................................... 192
recog_data_operand ............................... 517
recognizing insns ................................ 291
RECORD_TYPE ....................................... 179, 212
redirect_edge_and_branch, redirect_jump .... 284
redirect_edge_and_branch, redirect_jump ........ 286
reduc_smax_m instruction pattern ............... 336
reduc_smin_m instruction pattern ............... 336
reduc_splus_m instruction pattern ............... 336
reduc_umax_m instruction pattern ............... 336
reduc统战_m instruction pattern ............... 336
reduc_uplus_m instruction pattern ............... 337
reference ........................................... 179
REFERENCE_TYPE ................................... 179
reg .................................................. 144
reg and ‘/f’ ......................................... 133
reg and ‘/i’ ......................................... 133
reg and ‘/v’ ......................................... 133
reg, RTL sharing ................................... 176
REG_ALLOC_ORDER .................................. 420
REG_BR_PRED ......................................... 174
REG_BR_PROB ......................................... 174
REG_BR_PROB_BASE, BB_FREQ_BASE, count .... 284
REG_BR_PROB_BASE, EDGE_FREQUENCY .......... 284
REG_CC_SETTER .................................... 174
REG_CC_USER ...................................... 174
reg_class_contents ................................ 419
REG_CLASS_CONTENTS ............................... 426
REG_CLASS_FROM_CONSTRAINT .................... 434
REG_CLASS_FROM_LETTER ......................... 434
REG_CLASS_NAMES .................................. 426
REG_CROSSING_JUMP ............................... 172
REG_DEAD ............................................ 171
REG_DEAD, REG_UNUSED ............................. 287
REG_DEPANTI ....................................... 174
...
S
's' in constraint ........................................ 303
same_type_p ........................................... 181
Smode .................................................... 138
sat_fRACT ............................................. 158
satisfactm2 instruction pattern .................. 345
satisfactmn2 instruction pattern ............... 345
satisfies_constraint_m ......................... 332
SAVE_EXPR ............................................ 192
save_stack_block instruction pattern ....... 350
save_stack_function instruction pattern .... 350
save_stack_nonlocal instruction pattern .... 350
SBSS_SECTION_ASM_OP ............................... 492
Scalar evolutions ....................................... 273
scalars, returned as values ....................... 455
SCHED_GROUP_P ...................................... 134
Smode .................................................... 140
scratch .................................................. 149
scratch operands ...................................... 149
scratch, RTL sharing .................................. 176
scratch_operand ...................................... 298
SDATA_SECTION_ASM_OP ............................. 492
SDB_ALLOW_FORWARD_REFERENCES .............. 531
SDB_ALLOW_UNKNOWN_REFERENCES .............. 531
SDB_DEBUGGING_INFO ................................. 529
SDB_DELIM ............................................ 531
SDB_OUTPUT_SOURCE_LINE ............................ 531
SDmode ................................................ 138
sdot_prods instruction pattern .................. 337
search options ....................................... 367
SECONDARY_INPUT_RELOAD_CLASS ............. 431
SECONDARY_MEMORY_NEEDED ....................... 431
SECONDARY_MEMORY_NEEDED_MODE ............... 432
SECONDARY_MEMORY_NEEDED_RTX ................. 432
SECONDARY_OUTPUT_RELOAD_CLASS ............. 431
SECONDARY_RELOAD_CLASS ......................... 431
SELECT_CC_MODE ..................................... 478
sequence ............................................. 163
Sequence iterators .................................. 251
set .................................................. 159
set and '/f' .......................................... 133
SET_ASM_OP .......................................... 511
set_attr ............................................. 376
set_attr_alternative ............................... 376
set_bb_seq .......................................... 251
SET_BY_pieces_P .................................... 483
SET_DEST ............................................. 160
SET_IS_RETURN_P ..................................... 134
SET_LABEL_KIND ..................................... 168
set_optab_libfunc .................................. 468
SET_RATIO ............................................. 483
SET_SRC .............................................. 160
SET_TYPE_STRUCTURAL_EQUALITY .................. 179, 181
setmemm instruction pattern .................... 343
SETUP_FRAME_ADDRESSES ............................ 437
SF_SIZE ............................................. 415
SFmode ................................................ 138
sharing of RTL components ....................... 175
shift .................................................. 153
SHIFT_COUNT_TRUNCATED ......................... 544
SHLIB_SUFFIX ......................................... 516
SHORT_ACCUM_TYPE_SIZE ........................... 414
SHORT_FRACT_TYPE_SIZE ............................ 414
SHORT_IMMEDIATES_SIGN_EXTEND ................ 543
SHORT_TYPE_SIZE ................................... 413
sibcall_epilogue instruction pattern .......... 353
sibling call ......................................... 282
SIBLING_CALL_P ..................................... 134
SIG_ATOMIC_TYPE ..................................... 416
sign_extend .......................................... 157
sign_extract ......................................... 156
sign_extract, canonicalization of ............ 361
signed division ....................................... 152
signed division with signed saturation ... 152
signed maximum ...................................... 152
signed minimum ...................................... 152
SIMode ............................................... 138
simple constraints .................................. 302
sincos math function, implicit usage ........ 469
sinm2 instruction pattern ....................... 340
SIZE_ASM_OP ......................................... 505
SIZE_TYPE ............................................. 415
skip .................................................. 572
SLOW_BYTE_ACCESS .................................. 482
SLOW_UNALIGNED_ACCESS ............................ 482
smax ................................................. 152
Tuple specific accessors ............................................ 233

tuples ........................................................................ 222
type ......................................................................... 179
type declaration ......................................................... 184
TYPE_ALIGN .................................................. 179, 180, 209, 210
TYPE_ARG_TYPES ........................................... 179, 209
TYPE_ASM_OP ...................................................... 506
TYPE_ATTRIBUTES .................................................. 189
TYPE_BINFO ....................................................... 212
TYPE_BUILT_IN .................................................. 210
TYPE_CANONICAL .................................................. 179, 180
TYPE_CONTEXT ....................................................... 179, 209
TYPE_DECL .......................................................... 184
TYPE_FIELDS .................................................. 179, 209, 212
TYPE_HAS_ARRAY_NEW_OPERATOR .................................. 214
TYPE_HAS_DEFAULT_CONSTRUCTOR ................................ 214
TYPE_HAS_MUTABLE_P .............................................. 214
TYPE_HAS_NEW_OPERATOR ........................................ 214
TYPE_MAIN_VARIANT .................................. 179, 180, 209
TYPE_MAX_VALUE ................................................. 179
TYPE_METHOD_BASETYPE .................................. 179, 209
TYPE_METHODS .................................................. 212
TYPE_MIN_VALUE ................................................. 179
TYPE_NAME .................................................. 179, 180, 209, 210
TYPE_NOTHROW_P .................................................. 217
TYPE_OFFSET_BASETYPE .................................. 179, 209
TYPE_OPERAND_FMT .................................................. 506
TYPE_OVERLOADS_ARRAY_REF .................................. 214
TYPE_OVERLOADS_ARROW ...................................... 214
TYPE_OVERLOADS_CALL_EXPR .................................. 214
TYPE_POLYMORPHIC_P .............................................. 214
TYPE_PRECISION .................................................. 179, 209
TYPE_PTR .............................................................. 211
TYPE_PTRFN_P .................................................... 211
TYPE_PTRMEM_P .................................................... 209, 210
TYPE_PTRROB_P .................................................... 211
TYPE_PTRROBV_P ..................................................... 209
TYPE_QUAL ................................................................ 179, 209
TYPE_QUAL_CONST .................................................. 179, 209
TYPE_QUAL_RESTRICT ............................................. 179, 209
TYPE_QUAL_VOLATILE ............................................. 179, 209
TYPE_RAISES_EXCEPTIONS ........................................... 216
TYPE_SIZE .................................................. 179, 180, 209, 210
TYPE_STRUCTURAL_EQUALITY_P .................................. 179, 181
TYPE_UNQUALIFIED .................................................. 179, 209
TYPE_VFRIEND .................................................. 212
TYPE_NAME .................................................. 209
TYPE_NAME_TYPE .................................................. 209
TYPEOF_TYPE ....................................................... 209

U
UDmode ............................................................... 139
udiv ................................................................. 152
udivx3 instruction pattern .......................................... 336
udivx4 instruction pattern .......................................... 339
udot_prod instruction pattern ...................................... 337
UDQmode .............................................................. 139
UHAmode .............................................................. 139
<table>
<thead>
<tr>
<th>Instruction Pattern</th>
<th>Line</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>unsigned_sat_fract</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_fract_convert</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_float</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_fix</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_minimum</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_maximum</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_less</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_greater</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_division</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNORDERED_EXPR</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNKNOWN_TYPE</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unchanging</code>, in <code>call_insn</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unchanging</code>, in <code>jump_insn</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unchanging</code>, in <code>call_insn</code> and <code>insn</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unchanging</code>, in <code>mem</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNCHANGING</code>, in <code>subreg</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNCHANGING</code>, in <code>symbol_ref</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNEQ_EXPR</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNEQ_EXPR</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNGT_EXPR</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNION_TYPE</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNITS_PERWORD</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNKNOWN_TYPE</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNLIKEEXECUTEDTEXTSECTONAME</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNLT_EXPR</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UNORDERED_EXPR</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unshare_all rtl</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned division</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned division with unsigned saturation</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned greater than</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned less than</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned minimum and maximum</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_fix</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_float</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_fract_convert</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_sat_fract</code></td>
<td>132</td>
</tr>
<tr>
<td><code>unsigned_return</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UPDATE_PATH_HOST_CANONICALIZE</code></td>
<td>132</td>
</tr>
<tr>
<td><code>update_ssa</code></td>
<td>132</td>
</tr>
<tr>
<td><code>update_stmt</code></td>
<td>132</td>
</tr>
<tr>
<td><code>update_stmt_if_modified</code></td>
<td>132</td>
</tr>
<tr>
<td><code>UQMODE</code></td>
<td>132</td>
</tr>
<tr>
<td><code>usAshift</code></td>
<td>132</td>
</tr>
<tr>
<td><code>usMinus</code></td>
<td>132</td>
</tr>
<tr>
<td><code>usMult</code></td>
<td>132</td>
</tr>
<tr>
<td><code>usNeg</code></td>
<td>132</td>
</tr>
<tr>
<td><code>usPlus</code></td>
<td>132</td>
</tr>
<tr>
<td><code>usTruncate</code></td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd3</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>USAMode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd3</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usDiv3</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>use</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_CALLOC</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_LDASNEEDED</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_LOADPOSTDECREMENT</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_LOADPOST_INCREMENT</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_LOADPREDECREMENT</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_LOADPRE_INCREMENT</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_PARAM</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_PARAM</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_PARAM</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_SELECTSECTIONFORFUNCTIONS</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_STOREPOSTDECREMENT</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_STOREPOST_INCREMENT</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_STOREPREDECREMENT</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USE_STOREPRE_INCREMENT</code></td>
<td>132</td>
</tr>
<tr>
<td><code>used</code></td>
<td>132</td>
</tr>
<tr>
<td><code>used_in_symbol_ref</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USERLABELPREFIX</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USING_STMT</code></td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>usAdd4</code> instruction pattern</td>
<td>132</td>
</tr>
<tr>
<td><code>USAMode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USAMode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USAMode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USAMode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
<tr>
<td><code>USEQmode</code></td>
<td>132</td>
</tr>
</tbody>
</table>
Variable Location Debug Information in RTL .......................... 165
variable_size ............................................. 575
vashl3 instruction pattern ................................ 340
vashrs3 instruction pattern ................................ 340
vec_concat ................................................. 156
vec_duplicate ............................................. 157
VEC_EXTRACT EVEN_EXPR .................................. 199
vec_extract_evenm instruction pattern .................... 335
VEC_EXTRACT ODD_EXPR .................................. 199
vec_extract_oddm instruction pattern ..................... 335
vec_initm instruction pattern ................................ 335
VEC_INTERLEAVE_HIGH_EXPR ................................ 199
vec_interleave_highm instruction pattern ................. 335
VEC_INTERLEAVE_LOW_EXPR ................................ 199
vec_interleave_lowm instruction pattern ................. 335
VEC_LSHIFT_EXPR .......................................... 199
vec_merge .................................................. 156
VEC_PACK_FIX_TRUNC_EXPR ................................ 199
VEC_PACK_SAT_EXPR ........................................ 199
vec_pack_fix_trunc_m instruction pattern ............... 337
vec_pack_sat_m instruction pattern ....................... 337
VEC_PACK_TRUNC_EXPR ...................................... 199
vec_pack_trunc_m instruction pattern .................... 337
vec_pack_ufix_trunc_m instruction pattern ............... 337
vec_pack_usat_m instruction pattern ...................... 337
VEC_RSHIFT_EXPR .......................................... 199
vec_select .................................................. 156
vec_setz instruction pattern ................................ 335
vec_shl_m instruction pattern ................................ 337
vec_shr_m instruction pattern ................................ 337
VEC_UNPACK_FLOAT_HI_EXPR ................................ 199
VEC_UNPACK_FLOAT_LO_EXPR ................................ 199
VEC_UNPACK_HI_EXPR ........................................ 199
VEC_UNPACK_LO_EXPR ....................................... 199
vec_unpacks_float_hi_m instruction pattern .......... 338
vec_unpacks_float_lo_m instruction pattern .......... 338
vec_unpacks_hi_m instruction pattern .................... 337
vec_unpacks_lo_m instruction pattern .................... 337
vec_unpacku_float_hi_m instruction pattern ......... 338
vec_unpacku_float_lo_m instruction pattern ......... 338
vec_unpacku_hi_m instruction pattern .................... 338
vec_unpacku_lo_m instruction pattern .................... 338
VEC_WIDEN_MULT_HI_EXPR ................................ 199
VEC_WIDEN_MULT_LO_EXPR ................................ 199
vec_widen_smult_hi_m instruction pattern ............. 338
vec_widen_smult_lo_m instruction pattern ............. 338
vec_widen_umult_hi_m instruction pattern ............ 338
vec_widen_umult_lo_m instruction pattern ............ 338
vec .................................................. 179
vec operations ............................................ 156
VECTOR_CST ................................................. 189
VECTOR_STORE_FLAG_VALUE ................................ 547
virtual operands ............................................ 257
VIRTUAL_INCOMING_ARGS_REGNUM ......................... 145
VIRTUAL_OUTGOING_ARGS_REGNUM ......................... 146
VIRTUAL_STACK_DYNAMIC_REGNUM ......................... 145
VIRTUAL_STACK_VARS_REGNUM ............................ 145
VLIW .................................................. 380, 383
vlshr3 instruction pattern ................................ 340
VMS .................................................. 560
VMS_DEBUGGING_INFO ..................................... 531
VOID_TYPE .................................................. 179
VOIDmode ................................................... 140
volatile .................................................... 137
volatile, in insn, call_insn, jump_insn, code_label, barrier, and note .................. 132
volatile, in label_ref and reg_label .................... 132
volatile, in mem, asm_operands, and asm_input ........ 133
volatile, in reg ............................................. 133
volatile, in subreg ........................................ 134, 135
volatile, in symbol_ref .................................... 135
volatile memory references ................................ 137
volatile, in prefetch ...................................... 135
volatile, in prefetch, voting between constraint alternatives .................. 307
vrotl3 instruction pattern ................................ 340
vrotm3 instruction pattern ................................ 340
W
walk_dominator_tree ........................................ 266
walk_gimple_op ............................................. 255
walk_gimple_seq ............................................ 255
walk_gimplestmt ......................................... 254
walk_use_def_chains ...................................... 266
WCHAR_TYPE ............................................... 416
WCHAR_TYPE_SIZE .......................................... 416
which_alternative ........................................ 296
WHILE_BODY ................................................. 217
WHILE_COND ............................................... 217
WHILE_STMT ............................................... 217
whopr ...................................................... 587
WIDEST_HARDWARE_FP_SIZE ................................ 415
WINT_TYPE .................................................. 416
word_mode ................................................... 142
WORD_REGISTER_OPERATIONS ............................... 543
WORDS_BIG_ENDIAN .......................................... 404
WORDS_BIG_ENDIAN, effect on subreg ..................... 148
wpa .................................................. 587
X
'x' in constraint ........................................... 303
'x-host' ................................................... 565
XMode ...................................................... 140
XCOFF_DEBUGGING_INFO .................................. 526
XEXP .................................................. 128
XF_DEBUGGING_INFO ...................................... 526
XSIZE .................................................... 415
XMode ...................................................... 138
XINT ..................................................... 128
| 'xm-machine.h'                      | 560, 561 |
| xor                                 | 153      |
| xor, canonicalization of            | 361      |
| xor=x instruction pattern           | 336      |
| XSTR                                | 128      |
| XVEC                                | 129      |
| XVECEXP                             | 129      |
| XVECLEN                             | 129      |
| XWINT                               | 128      |
| zero_extend                         | 157      |
| zero_extend=x2 instruction pattern  | 345      |
| zero_extract                        | 156      |
| zero_extract, canonicalization of    | 361      |