Using the GNU Compiler Collection

Richard M. Stallman

Last updated 21 May 2002

for MIPS SDE v5.0 / gcc-2.96

For MIPS SDE v5.0 / GCC Version 2.96

Published by the Free Software Foundation
59 Temple Place - Suite 330
Boston, MA 02111-1307, USA
Last printed April, 1998.
Printed copies are available for $50 each.
ISBN 1-882114-37-X

Introduction

This manual documents how to run the GNU compiler, as well as its new features and incompatibilities, and how to report bugs. It corresponds to GCC version 2.96.

1 Compile C, C++, Objective C, Fortran, Java or CHILL

Several versions of the compiler (C, C++, Objective C, Fortran, Java and CHILL) are integrated; this is why we use the name "GNU Compiler Collection". GCC can compile programs written in any of these languages. The Fortran and CHILL compilers are described in separate manuals. The Java compiler currently has no manual documenting it.

"GCC" is a common shorthand term for the GNU Compiler Collection. This is both the most general name for the compiler, and the name used when the emphasis is on compiling C programs (as the abbreviation formerly stood for "GNU C Compiler").

When referring to C++ compilation, it is usual to call the compiler "G++". Since there is only one compiler, it is also accurate to call it "GCC" no matter what the language context; however, the term "G++" is more useful when the emphasis is on compiling C++ programs.

We use the name "GCC" to refer to the compilation system as a whole, and more specifically to the language-independent part of the compiler. For example, we refer to the optimization options as affecting the behavior of "GCC" or sometimes just "the compiler".

Front ends for other languages, such as Ada 95 and Pascal exist but have not yet been integrated into GCC. These front-ends, like that for C++, are built in subdirectories of GCC and link to it. The result is an integrated compiler that can compile programs written in C, C++, Objective C, or any of the languages for which you have installed front ends.

In this manual, we only discuss the options for the C, Objective-C, and C++ compilers and those of the GCC core. Consult the documentation of the other front ends for the options to use when compiling programs written in other languages.

G++ is a *compiler*, not merely a preprocessor. G++ builds object code directly from your C++ program source. There is no intermediate C version of the program. (By contrast, for example, some other implementations use a program that generates a C program from your C++ source.) Avoiding an intermediate C representation of the program means that you get better object code, and better debugging information. The GNU debugger, GDB, works with this information in the object code to give you comprehensive C++ source-level editing capabilities (see section "C and C++" in *Debugging with GDB*).

2 GCC Command Options

When you invoke GCC, it normally does preprocessing, compilation, assembly and linking. The "overall options" allow you to stop this process at an intermediate stage. For example, the '`-c`' option says not to run the linker. Then the output consists of object files output by the assembler.

Other options are passed on to one stage of processing. Some options control the preprocessor and others the compiler itself. Yet other options control the assembler and linker; most of these are not documented here, since you rarely need to use any of them.

Most of the command line options that you can use with GCC are useful for C programs; when an option is only useful with another language (usually C++), the explanation says so explicitly. If the description for a particular option does not mention a source language, you can use that option with all supported languages.

See Section 2.3 [Compiling C++ Programs], page 10, for a summary of special options for compiling C++ programs.

The `gcc` program accepts options and file names as operands. Many options have multiletter names; therefore multiple single-letter options may *not* be grouped: '`-dr`' is very different from '`-d -r`'.

You can mix options and other arguments. For the most part, the order you use doesn't matter. Order does matter when you use several options of the same kind; for example, if you specify '`-L`' more than once, the directories are searched in the order specified.

Many options have long names starting with '`-f`' or with '`-W`'—for example, '`-fforce-mem`', '`-fstrength-reduce`', '`-Wformat`' and so on. Most of these have both positive and negative forms; the negative form of '`-ffoo`' would be '`-fno-foo`'. This manual documents only one of these two forms, whichever one is not the default.

2.1 Option Summary

Here is a summary of all the options, grouped by type. Explanations are in the following sections.

Overall Options

See Section 2.2 [Options Controlling the Kind of Output], page 8.

`-c -S -E -o` *file* `-pipe -pass-exit-codes -v --help -x` *language*

C Language Options

See Section 2.4 [Options Controlling C Dialect], page 11.

```
-ansi -fstd -fallow-single-precision -fcond-mismatch -fno-asm
-fno-builtin -ffreestanding -fhosted -fsigned-bitfields
-funsigned-bitfields -fsigned-char -funsigned-char
-fwritable-strings -traditional -traditional-cpp -trigraphs
```

C++ Language Options

See Section 2.5 [Options Controlling C++ Dialect], page 15.

```
-fno-access-control -fcheck-new -fconserve-space
-fdollars-in-identifiers -fno-elide-constructors -fexternal-templates
-ffor-scope -fno-for-scope -fno-gnu-keywords -fguiding-decls -fhonor-std
-fhuge-objects -fno-implicit-templates -finit-priority
-fno-implement-inlines -fname-mangling-version-n
```

`-fno-default-inline -fno-operator-names -fno-optional-diags -fpermissive`
`-frepo -fstrict-prototype -fsquangle -ftemplate-depth-`*n*
`-fuse-cxa-atexit -fvtable-thunks -nostdinc++ -Wctor-dtor-privacy`
`-Wno-deprecated -Weffc++ -Wno-non-template-friend -Wnon-virtual-dtor`
`-Wold-style-cast -Woverloaded-virtual -Wno-pmf-conversions -Wreorder`
`-Wsign-promo -Wsynth`

Warning Options

See Section 2.6 [Options to Request or Suppress Warnings], page 21.

`-fsyntax-only -pedantic -pedantic-errors`
`-w -W -Wall -Waggregate-return`
`-Wcast-align -Wcast-qual -Wchar-subscripts -Wcomment`
`-Wconversion -Werror -Wformat`
`-Wid-clash-`*len* `-Wimplicit -Wimplicit-int`
`-Wimplicit-function-declaration -Wimport`
`-Werror-implicit-function-declaration -Wfloat-equal -Winline`
`-Wlarger-than-`*len* `-Wlong-long`
`-Wmain -Wmissing-declarations -Wmissing-noreturn`
`-Wparentheses -Wparentheses-else -Wpointer-arith -Wredundant-decls`
`-Wreturn-type -Wlongjmp-clobbers -Wshadow -Wsign-compare`
`-Wstrict-prototypes -Wswitch -Wtraditional`
`-Wmultichar -Wno-import -Wpacked -Wpadded`
`-Wparentheses -Wpointer-arith -Wredundant-decls`
`-Wreturn-type -Wshadow -Wsign-compare -Wswitch`
`-Wtrigraphs -Wundef -Wuninitialized -Wunknown-pragmas -Wunreachable-code`
`-Wunused -Wunused-function -Wunused-label -Wunused-parameter`
`-Wunused-variable -Wunused-value -Wwrite-strings`

C-only Warning Options

`-Wbad-function-cast -Wmissing-prototypes -Wnested-externs`
`-Wstrict-prototypes -Wtraditional`

Debugging Options

See Section 2.7 [Options for Debugging Your Program or GCC], page 29.

`-d`*letters* `-fdump-unnumbered -fdump-translation-unit-`*file*
`-fpretend-float -fprofile-arcs -ftest-coverage`
`-g -g`*level* `-gcoff -gdwarf -gdwarf-1 -gdwarf-1+ -gdwarf-2`
`-ggdb -gstabs -gstabs+ -gxcoff -gxcoff+`
`-p -pg -print-file-name=`*library* `-print-libgcc-file-name`
`-print-prog-name=`*program* `-print-search-dirs -save-temps -time`

Optimization Options

See Section 2.8 [Options that Control Optimization], page 33.

`-falign-functions=`*n* `-falign-labels=`*n* `-falign-loops=`*n*
`-falign-jumps=`*n* `-fbranch-probabilities`
`-fcaller-saves -fcse-follow-jumps -fcse-skip-blocks`
`-flive-range`
`-fdelayed-branch -fdelete-null-pointer-checks -fexpensive-optimizations`
`-ffast-math -ffloat-store -fforce-addr -fforce-mem -fno-math-errno`
`-fdata-sections -ffunction-sections -fgcse`
`-finline-functions -finline-limit=`*n* `-fkeep-inline-functions`

```
-fmove-all-movables  -fno-default-inline  -fno-defer-pop
-fno-function-cse  -fno-inline  -fno-peephole
-fomit-frame-pointer  -foptimize-register-moves -foptimize-sibling-calls
-fregmove -frerun-cse-after-loop  -frerun-loop-opt  -freduce-all-givs
-fschedule-insns  -fschedule-insns2  -fssa -fstrength-reduce
-fstrict-aliasing  -fthread-jumps  -funroll-all-loops
-funroll-loops
-O  -O0  -O1  -O2  -O3 -Os
```

Preprocessor Options

See Section 2.9 [Options Controlling the Preprocessor], page 40.

```
-Aquestion(answer)  -C  -dD  -dM  -dN
-Dmacro[=defn]  -E  -H
-idirafter dir
-include file  -imacros file
-iprefix file  -iwithprefix dir
-iwithprefixbefore dir  -isystem dir -isystem-c++ dir
-M  -MD  -MM  -MMD  -MG  -nostdinc  -P  -trigraphs
-undef  -Umacro  -Wp,option
```

Assembler Option

See Section 2.10 [Passing Options to the Assembler], page 43.

```
-Wa,option
```

Linker Options

See Section 2.11 [Options for Linking], page 43.

```
object-file-name  -llibrary
-nostartfiles  -nodefaultlibs  -nostdlib
-s  -static  -shared  -symbolic
-Wl,option  -Xlinker option
-u symbol
```

Directory Options

See Section 2.12 [Options for Directory Search], page 46.

```
-Bprefix  -Idir  -I-  -Ldir  -specs=file
```

Options specific to MIPS Processors and/or MIPS SDE

See Section 2.15 [Different CPUs and Configurations], page 52

```
-mcpu=cpu type
-mips1  -mips2  -mips3  -mips4  -mips5
-mips32  -mips32r2 -mips64 -mips64r2
-mips16  -mips16e  -msmartmips  -mips3D
-mcode-xonly  -mno-data-in-code
-muse-all-regs -mbranch-likely
-mcheck-zero-division  -mno-check-zero-division
-mdiv-checks  -mno-div-checks
-mconst-mult  -mno-const-mult
-mfp32  -mfp64  -mgp32  -mgp64
-msplit-addresses  -mrnames  -mgpopt
-mstats
```

```
-mmemcpy
-msoft-float  -mhard-float  -msingle-float
-mslow-mul  -mno-mul  -mmad  -mno-mad
-membedded-data  -mno-embedded-data  -mgpconst  -mno-gpconst
-muninit-const-in-rodata  -mno-uninit-const-in-rodata
-mlong-calls  -mcommon-prolog  -mentry
-EL  -EB
-G num
-nocpp

-mabi=32  -mabi=n32  -mabi=64  -mabi=eabi  -mabi=meabi
-mabicalls  -mno-abicalls
-mint64  -mlong64  -mgas  -mmips-as
-mhalf-pic  -mno-half-pic  -membedded-pic  -mno-embedded-pic
-mmips-tfile  -mno-mips-tfile
```

Code Generation Options

See Section 2.16 [Options for Code Generation Conventions], page 61.

```
-fcall-saved-reg  -fcall-used-reg
-fexceptions  -funwind-tables  -ffixed-reg  -finhibit-size-directive
-fcheck-memory-usage  -fprefix-function-name
-fno-common  -fcommon  -fno-ident  -fno-gnu-linker
-fpcc-struct-return  -fpic  -fPIC
-freg-struct-return  -fshared-data  -fshort-enums
-fshort-double  -fvolatile  -fvolatile-global -fvolatile-static
-funaligned-pointers  -funaligned-struct-hack
-foptimize-comparisons
-fverbose-asm  -fpack-struct  -fstack-check
-fstack-limit-register=reg  -fstack-limit-symbol=sym
-fargument-alias  -fargument-noalias
-fargument-noalias-global
-fleading-underscore
```

2.2 Options Controlling the Kind of Output

Compilation can involve up to four stages: preprocessing, compilation proper, assembly and linking, always in that order. The first three stages apply to an individual source file, and end by producing an object file; linking combines all the object files (those newly compiled, and those specified as input) into an executable file.

For any given input file, the file name suffix determines what kind of compilation is done:

file`.c` C source code which must be preprocessed.

file`.i` C source code which should not be preprocessed.

file`.ii` C++ source code which should not be preprocessed.

file`.m` Objective-C source code. Note that you must link with the library ‘`libobjc.a`’ to make an Objective-C program work.

file`.h` C header file (not to be compiled or linked).

file`.cc`
file`.cxx`
file`.cpp`
file`.C` C++ source code which must be preprocessed. Note that in '`.cxx`', the last two letters must both be literally '`x`'. Likewise, '`.C`' refers to a literal capital C.

file`.s` Assembler code.

file`.S`
file`.sx` Assembler code which must be preprocessed. The *file*.sx form is used in Windows, where filenames are case-insensitive.

other An object file to be fed straight into linking. Any file name with no recognized suffix is treated this way.

You can specify the input language explicitly with the '`-x`' option:

`-x` *language*
Specify explicitly the *language* for the following input files (rather than letting the compiler choose a default based on the file name suffix). This option applies to all following input files until the next '`-x`' option. Possible values for *language* are:

```
c  objective-c  c++
c-header  cpp-output  c++-cpp-output
assembler  assembler-with-cpp
```

`-x none` Turn off any specification of a language, so that subsequent files are handled according to their file name suffixes (as they are if '`-x`' has not been used at all).

`-pass-exit-codes`
Normally the `gcc` program will exit with the code of 1 if any phase of the compiler returns a non-success return code. If you specify '`-pass-exit-codes`', the `gcc` program will instead return with numerically highest error produced by any phase that returned an error indication.

If you only want some of the stages of compilation, you can use '`-x`' (or filename suffixes) to tell `gcc` where to start, and one of the options '`-c`', '`-S`', or '`-E`' to say where `gcc` is to stop. Note that some combinations (for example, '`-x cpp-output -E`' instruct `gcc` to do nothing at all.

`-c` Compile or assemble the source files, but do not link. The linking stage simply is not done. The ultimate output is in the form of an object file for each source file.

By default, the object file name for a source file is made by replacing the suffix '`.c`', '`.i`', '`.s`', etc., with '`.o`'.

Unrecognized input files, not requiring compilation or assembly, are ignored.

`-S` Stop after the stage of compilation proper; do not assemble. The output is in the form of an assembler code file for each non-assembler input file specified.

By default, the assembler file name for a source file is made by replacing the suffix '`.c`', '`.i`', etc., with '`.s`'.

Input files that don't require compilation are ignored.

`-E` Stop after the preprocessing stage; do not run the compiler proper. The output is in the form of preprocessed source code, which is sent to the standard output.

Input files which don't require preprocessing are ignored.

`-o` *file* Place output in file *file*. This applies regardless to whatever sort of output is being produced, whether it be an executable file, an object file, an assembler file or preprocessed C code.

Since only one output file can be specified, it does not make sense to use '`-o`' when compiling more than one input file, unless you are producing an executable file as output.

If '`-o`' is not specified, the default is to put an executable file in '`a.out`', the object file for '*source*`.`*suffix*' in '*source*`.o`', its assembler file in '*source*`.s`', and all preprocessed C source on standard output.

`-v` Print (on standard error output) the commands executed to run the stages of compilation. Also print the version number of the compiler driver program and of the preprocessor and the compiler proper.

`-pipe` Use pipes rather than temporary files for communication between the various stages of compilation. This fails to work on some systems where the assembler is unable to read from a pipe; but the GNU assembler has no trouble.

`--print-multi-directory`
Prints the name of the library "multilib" subdirectory that will be used for this set of options.

`--print-search-dirs`
Prints the list of directories searched for programs (eg '`cc1`', '`cpp`', etc) and libraries.

`--print-file-name=`*file*
`--print-prog-name=`*program*
Prints the location of the library or program file given the previous options. For example '`--print-file-name=cc1`' will tell you which program is being used for the main compiler pass, and '`--print-file-name=libc.a`' will tell you where the C library is coming from.

`--help` Print (on the standard output) a description of the command line options understood by `gcc`. If the `-v` option is also specified then `--help` will also be passed on to the various processes invoked by `gcc`, so that they can display the command line options they accept. If the `-W` option is also specified then command line options which have no documentation associated with them will also be displayed.

2.3 Compiling C++ Programs

C++ source files conventionally use one of the suffixes '`.C`', '`.cc`', '`.cpp`', '`.c++`', '`.cp`', or '`.cxx`'; preprocessed C++ files use the suffix '`.ii`'. GCC recognizes files with these names and compiles them as C++ programs even if you call the compiler the same way as for compiling C programs (usually with the name `gcc`).

However, C++ programs often require class libraries as well as a compiler that understands the C++ language—and under some circumstances, you might want to compile programs from standard input, or otherwise without a suffix that flags them as C++ programs. `g++` is a program that calls GCC with the default language set to C++, and automatically specifies linking against the C++ library. On many systems, the script `g++` is also installed with the name `c++`.

When you compile C++ programs, you may specify many of the same command-line options that you use for compiling programs in any language; or command-line options meaningful for C and related languages; or options that are meaningful only for C++ programs. See Section 2.4 [Options Controlling C Dialect], page 11, for explanations of options for languages related to C. See Section 2.5 [Options Controlling C++ Dialect], page 15, for explanations of options that are meaningful only for C++ programs.

2.4 Options Controlling C Dialect

The following options control the dialect of C (or languages derived from C, such as C++ and Objective C) that the compiler accepts:

`-ansi` In C mode, support all ANSI standard C programs. In C++ mode, remove GNU extensions that conflict with ISO C++.

This turns off certain features of GCC that are incompatible with ANSI C (when compiling C code), or of standard C++ (when compiling C++ code), such as the `asm` and `typeof` keywords, and predefined macros such as `unix` and `vax` that identify the type of system you are using. It also enables the undesirable and rarely used ANSI trigraph feature. For the C compiler, it disables recognition of C++ style '`//`' comments as well as the `inline` keyword.

The alternate keywords `__asm__`, `__extension__`, `__inline__` and `__typeof__` continue to work despite '`-ansi`'. You would not want to use them in an ANSI C program, of course, but it is useful to put them in header files that might be included in compilations done with '`-ansi`'. Alternate predefined macros such as `__unix__` and `__vax__` are also available, with or without '`-ansi`'.

The '`-ansi`' option does not cause non-ANSI programs to be rejected gratuitously. For that, '`-pedantic`' is required in addition to '`-ansi`'. See Section 2.6 [Warning Options], page 21.

The macro `__STRICT_ANSI__` is predefined when the '`-ansi`' option is used. Some header files may notice this macro and refrain from declaring certain functions or defining certain macros that the ANSI standard doesn't call for; this is to avoid interfering with any programs that might use these names for other things.

The functions `alloca`, `abort`, `exit`, and `_exit` are not builtin functions when '`-ansi`' is used.

`-fstd=` Determine the language standard. A value for this option must be provided; possible values are

- iso9899:1990 Same as -ansi
- iso9899:199409 ISO C as modified in amend. 1
- iso9899:199x ISO C 9x
- c89 same as -std=iso9899:1990
- c9x same as -std=iso9899:199x
- gnu89 default, iso9899:1990 + gnu extensions
- gnu9x iso9899:199x + gnu extensions

Even when this option is not specified, you can still use some of the features of newer standards in so far as they do not conflict with previous C standards. For example, you may use __restrict__ even when -fstd=c9x is not specified.

`-fno-asm` Do not recognize `asm`, `inline` or `typeof` as a keyword, so that code can use these words as identifiers. You can use the keywords `__asm__`, `__inline__` and `__typeof__` instead. '`-ansi`' implies '`-fno-asm`'.

In C++, this switch only affects the `typeof` keyword, since `asm` and `inline` are standard keywords. You may want to use the '`-fno-gnu-keywords`' flag instead, which has the same effect.

`-fno-builtin`

Don't recognize builtin functions that do not begin with '`__builtin_`' as prefix. Currently, the functions affected include `abort`, `abs`, `alloca`, `cos`, `cosf`, `cosl`, `exit`, `_exit`, `fabs`, `fabsf`, `fabsl`, `ffs`, `labs`, `memcmp`, `memcpy`, `memset`, `sin`, `sinf`, `sinl`, `sqrt`, `sqrtf`, `sqrtl`, `strcmp`, `strcpy`, and `strlen`.

GCC normally generates special code to handle certain builtin functions more efficiently; for instance, calls to `alloca` may become single instructions that adjust the stack directly, and calls to `memcpy` may become inline copy loops. The resulting code is often both smaller and faster, but since the function calls no longer appear as such, you cannot set a breakpoint on those calls, nor can you change the behavior of the functions by linking with a different library.

The '`-ansi`' option prevents `alloca`, `ffs` and `_exit` from being builtin functions, since these functions do not have an ANSI standard meaning.

`-fhosted`

Assert that compilation takes place in a hosted environment. This implies '`-fbuiltin`'. A hosted environment is one in which the entire standard library is available, and in which `main` has a return type of `int`. Examples are nearly everything except a kernel. This is equivalent to '`-fno-freestanding`'.

`-ffreestanding`

Assert that compilation takes place in a freestanding environment. This implies '`-fno-builtin`'. A freestanding environment is one in which the standard library may not exist, and program startup may not necessarily be at `main`. The most obvious example is an OS kernel. This is equivalent to '`-fno-hosted`'.

`-trigraphs`

Support ANSI C trigraphs. You don't want to know about this brain-damage. The '`-ansi`' option implies '`-trigraphs`'.

`-traditional`

Attempt to support some aspects of traditional C compilers. Specifically:

- All `extern` declarations take effect globally even if they are written inside of a function definition. This includes implicit declarations of functions.
- The newer keywords `typeof`, `inline`, `signed`, `const` and `volatile` are not recognized. (You can still use the alternative keywords such as `__typeof__`, `__inline__`, and so on.)
- Comparisons between pointers and integers are always allowed.

- Integer types `unsigned short` and `unsigned char` promote to `unsigned int`.
- Out-of-range floating point literals are not an error.
- Certain constructs which ANSI regards as a single invalid preprocessing number, such as '`0xe-0xd`', are treated as expressions instead.
- String "constants" are not necessarily constant; they are stored in writable space, and identical looking constants are allocated separately. (This is the same as the effect of '`-fwritable-strings`'.)
- All automatic variables not declared `register` are preserved by `longjmp`. Ordinarily, GNU C follows ANSI C: automatic variables not declared `volatile` may be clobbered.
- The character escape sequences '`\x`' and '`\a`' evaluate as the literal characters '`x`' and '`a`' respectively. Without '`-traditional`', '`\x`' is a prefix for the hexadecimal representation of a character, and '`\a`' produces a bell.

You may wish to use '`-fno-builtin`' as well as '`-traditional`' if your program uses names that are normally GNU C builtin functions for other purposes of its own.

You cannot use '`-traditional`' if you include any header files that rely on ANSI C features. Some vendors are starting to ship systems with ANSI C header files and you cannot use '`-traditional`' on such systems to compile files that include any system headers.

The '`-traditional`' option also enables '`-traditional-cpp`', which is described next.

`-traditional-cpp`

Attempt to support some aspects of traditional C preprocessors. Specifically:

- Comments convert to nothing at all, rather than to a space. This allows traditional token concatenation.
- In a preprocessing directive, the '`#`' symbol must appear as the first character of a line.
- Macro arguments are recognized within string constants in a macro definition (and their values are stringified, though without additional quote marks, when they appear in such a context). The preprocessor always considers a string constant to end at a newline.
- The predefined macro `__STDC__` is not defined when you use '`-traditional`', but `__GNUC__` is (since the GNU extensions which `__GNUC__` indicates are not affected by '`-traditional`'). If you need to write header files that work differently depending on whether '`-traditional`' is in use, by testing both of these predefined macros you can distinguish four situations: GNU C, traditional GNU C, other ANSI C compilers, and other old C compilers. The predefined macro `__STDC_VERSION__` is also not defined when you use '`-traditional`'. See section "Standard Predefined Macros" in *The C Preprocessor*, for more discussion of these and other predefined macros.
- The preprocessor considers a string constant to end at a newline (unless the newline is escaped with '`\`'). (Without '`-traditional`', string constants can contain the newline character as typed.)

`-fcond-mismatch`

Allow conditional expressions with mismatched types in the second and third arguments. The value of such an expression is void.

`-funsigned-char`

Let the type `char` be unsigned, like `unsigned char`.

Each kind of machine has a default for what `char` should be. It is either like `unsigned char` by default or like `signed char` by default.

Ideally, a portable program should always use `signed char` or `unsigned char` when it depends on the signedness of an object. But many programs have been written to use plain `char` and expect it to be signed, or expect it to be unsigned, depending on the machines they were written for. This option, and its inverse, let you make such a program work with the opposite default.

The type `char` is always a distinct type from each of `signed char` or `unsigned char`, even though its behavior is always just like one of those two.

`-fsigned-char`

Let the type `char` be signed, like `signed char`.

Note that this is equivalent to '`-fno-unsigned-char`', which is the negative form of '`-funsigned-char`'. Likewise, the option '`-fno-signed-char`' is equivalent to '`-funsigned-char`'.

You may wish to use '`-fno-builtin`' as well as '`-traditional`' if your program uses names that are normally GNU C builtin functions for other purposes of its own.

You cannot use '`-traditional`' if you include any header files that rely on ANSI C features. Some vendors are starting to ship systems with ANSI C header files and you cannot use '`-traditional`' on such systems to compile files that include any system headers.

`-fsigned-bitfields`
`-funsigned-bitfields`
`-fno-signed-bitfields`
`-fno-unsigned-bitfields`

These options control whether a bitfield is signed or unsigned, when the declaration does not use either `signed` or `unsigned`. By default, such a bitfield is signed, because this is consistent: the basic integer types such as `int` are signed types.

However, when '`-traditional`' is used, bitfields are all unsigned no matter what.

`-fwritable-strings`

Store string constants in the writable data segment and don't uniquize them. This is for compatibility with old programs which assume they can write into string constants. The option '`-traditional`' also has this effect.

Writing into string constants is a very bad idea; "constants" should be constant.

`-fallow-single-precision`

Do not promote single precision math operations to double precision, even when compiling with '`-traditional`'.

Traditional K&R C promotes all floating point operations to double precision, regardless of the sizes of the operands. On the architecture for which you are compiling,

single precision may be faster than double precision. If you must use '`-traditional`', but want to use single precision operations when the operands are single precision, use this option. This option has no effect when compiling with ANSI or GNU C conventions (the default).

`-fshort-wchar`
: Override the underlying type for '`wchar_t`' to be '`short unsigned int`' instead of the default for the target. This option is useful for building programs to run under WINE.

2.5 Options Controlling C++ Dialect

This section describes the command-line options that are only meaningful for C++ programs; but you can also use most of the GNU compiler options regardless of what language your program is in. For example, you might compile a file `firstClass.C` like this:

```
g++ -g -frepo -O -c firstClass.C
```

In this example, only '`-frepo`' is an option meant only for C++ programs; you can use the other options with any language supported by GCC.

Here is a list of options that are *only* for compiling C++ programs:

`-fno-access-control`
: Turn off all access checking. This switch is mainly useful for working around bugs in the access control code.

`-fcheck-new`
: Check that the pointer returned by `operator new` is non-null before attempting to modify the storage allocated. The current Working Paper requires that `operator new` never return a null pointer, so this check is normally unnecessary.

  An alternative to using this option is to specify that your `operator new` does not throw any exceptions; if you declare it '`throw()`', g++ will check the return value. See also '`new (nothrow)`'.

`-fconserve-space`
: Put uninitialized or runtime-initialized global variables into the common segment, as C does. This saves space in the executable at the cost of not diagnosing duplicate definitions. If you compile with this flag and your program mysteriously crashes after `main()` has completed, you may have an object that is being destroyed twice because two definitions were merged.

  This option is no longer useful on most targets, now that support has been added for putting variables into BSS without making them common.

`-fdollars-in-identifiers`
: Accept '`$`' in identifiers. You can also explicitly prohibit use of '`$`' with the option '`-fno-dollars-in-identifiers`'. (GNU C allows '`$`' by default on most target systems, but there are a few exceptions.) Traditional C allowed the character '`$`' to form part of identifiers. However, ANSI C and C++ forbid '`$`' in identifiers.

`-fembedded-cxx`
: In compliance with the Embedded C++ specification, make the use of templates, exception handling, multiple inheritance, or RTTI illegal. Attempts to use namespaces

are also not allowed. This makes the use of these keywords result in warnings by default: `template`, `typename`, `catch`, `throw`, `try`, `using`, `namespace`, `dynamic_cast`, `static_cast`, `reinterpret_cast`, `const_cast`, and `typeid`. To make the warnings for these things be given as errors, add the `-pedantic-errors` flag.

`-fno-elide-constructors`
The C++ standard allows an implementation to omit creating a temporary which is only used to initialize another object of the same type. Specifying this option disables that optimization, and forces g++ to call the copy constructor in all cases.

`-fno-enforce-eh-specs`
Don't check for violation of exception specifications at runtime. This option violates the C++ standard, but may be useful for reducing code size in production builds, much like defining '`NDEBUG`'. The compiler will still optimize based on the exception specifications.

`-fexternal-templates`
Cause template instantiations to obey '`#pragma interface`' and '`implementation`'; template instances are emitted or not according to the location of the template definition.
See Section 4.6 [Template Instantiation], page 120, for more information.

This option is deprecated.

`-falt-external-templates`
Similar to -fexternal-templates, but template instances are emitted or not according to the place where they are first instantiated.
See Section 4.6 [Template Instantiation], page 120, for more information.

This option is deprecated.

`-ffor-scope`
`-fno-for-scope`
If -ffor-scope is specified, the scope of variables declared in a *for-init-statement* is limited to the '`for`' loop itself, as specified by the C++ standard. If -fno-for-scope is specified, the scope of variables declared in a *for-init-statement* extends to the end of the enclosing scope, as was the case in old versions of gcc, and other (traditional) implementations of C++.

The default if neither flag is given to follow the standard, but to allow and give a warning for old-style code that would otherwise be invalid, or have different behavior.

`-fno-gnu-keywords`
Do not recognize `typeof` as a keyword, so that code can use this word as an identifier. You can use the keyword `__typeof__` instead. '`-ansi`' implies '`-fno-gnu-keywords`'.

`-fguiding-decls`
Treat a function declaration with the same type as a potential function template instantiation as though it declares that instantiation, not a normal function. If a definition is given for the function later in the translation unit (or another translation unit if the target supports weak symbols), that definition will be used; otherwise the template will be instantiated. This behavior reflects the C++ language prior to September 1996, when guiding declarations were removed.

This option implies '`-fname-mangling-version-0`', and will not work with other name mangling versions. Like all options that change the ABI, all C++ code, *including libgcc.a* must be built with the same setting of this option.

`-fno-implicit-templates`
: Never emit code for templates which are instantiated implicitly (i.e. by use); only emit code for explicit instantiations. See Section 4.6 [Template Instantiation], page 120, for more information.

`-fhonor-std`
: Treat the `namespace std` as a namespace, instead of ignoring it. For compatibility with earlier versions of g++, the compiler will, by default, ignore `namespace-declarations`, `using-declarations`, `using-directives`, and `namespace-names`, if they involve `std`.

`-fhuge-objects`
: Support virtual function calls for objects that exceed the size representable by a '`short int`'. Users should not use this flag by default; if you need to use it, the compiler will tell you so.

  This flag is not useful when compiling with -fvtable-thunks.

  Like all options that change the ABI, all C++ code, *including libgcc* must be built with the same setting of this option.

`-fmessage-length=`*n*
: Try to format error messages so that they fit on lines of about *n* characters. The default is 72 characters. If *n* is zero, then no line-wrapping will be done; each error message will appear on a single line.

`-fno-implicit-templates`
: Never emit code for non-inline templates which are instantiated implicitly (i.e. by use); only emit code for explicit instantiations. See Section 4.6 [Template Instantiation], page 120, for more information.

`-fno-implicit-inline-templates`
: Don't emit code for implicit instantiations of inline templates, either. The default is to handle inlines differently so that compiles with and without optimization will need the same set of explicit instantiations.

`-finit-priority`
: Support '`__attribute__ ((init_priority (n)))`' for controlling the order of initialization of file-scope objects. On ELF targets, this requires GNU ld 2.10 or later.

`-fno-implement-inlines`
: To save space, do not emit out-of-line copies of inline functions controlled by '`#pragma implementation`'. This will cause linker errors if these functions are not inlined everywhere they are called.

`-fms-extensions`
: Disable pedwarns about constructs used in MFC, such as implicit int and getting a pointer to member function via non-standard syntax.

`-fname-mangling-version-`*n*
: Control the way in which names are mangled. Version 0 is compatible with versions of g++ before 2.8. Version 1 is the default. Version 1 will allow correct mangling of

function templates. For example, version 0 mangling does not mangle foo<int, double> and foo<int, char> given this declaration:

```
template <class T, class U> void foo(T t);
```

Like all options that change the ABI, all C++ code, *including libgcc* must be built with the same setting of this option.

`-fno-operator-names`
: Do not treat the operator name keywords `and`, `bitand`, `bitor`, `compl`, `not`, `or` and `xor` as synonyms as keywords.

`-fno-optional-diags`
: Disable diagnostics that the standard says a compiler does not need to issue. Currently, the only such diagnostic issued by g++ is the one for a name having multiple meanings within a class.

`-fpermissive`
: Downgrade messages about nonconformant code from errors to warnings. By default, g++ effectively sets '`-pedantic-errors`' without '`-pedantic`'; this option reverses that. This behavior and this option are superseded by '`-pedantic`', which works as it does for GNU C.

`-frepo`
: Enable automatic template instantiation. This option also implies '`-fno-implicit-templates`'. See Section 4.6 [Template Instantiation], page 120, for more information.

`-fno-rtti`
: Disable generation of information about every class with virtual functions for use by the C++ runtime type identification features ('`dynamic_cast`' and '`typeid`'). If you don't use those parts of the language, you can save some space by using this flag. Note that exception handling uses the same information, but it will generate it as needed.

`-fstrict-prototype`
: Within an '`extern "C"`' linkage specification, treat a function declaration with no arguments, such as '`int foo ();`', as declaring the function to take no arguments. Normally, such a declaration means that the function `foo` can take any combination of arguments, as in C. '`-pedantic`' implies '`-fstrict-prototype`' unless overridden with '`-fno-strict-prototype`'.

  Specifying this option will also suppress implicit declarations of functions.

  This flag no longer affects declarations with C++ linkage.

`-fsquangle`
`-fno-squangle`
: '`-fsquangle`' will enable a compressed form of name mangling for identifiers. In particular, it helps to shorten very long names by recognizing types and class names which occur more than once, replacing them with special short ID codes. This option also requires any C++ libraries being used to be compiled with this option as well. The compiler has this disabled (the equivalent of '`-fno-squangle`') by default.

  Like all options that change the ABI, all C++ code, *including libgcc.a* must be built with the same setting of this option.

`-ftemplate-depth-`*n*

Set the maximum instantiation depth for template classes to *n*. A limit on the template instantiation depth is needed to detect endless recursions during template class instantiation. ANSI/ISO C++ conforming programs must not rely on a maximum depth greater than 17.

`-fuse-cxa-atexit`

Register destructors for objects with static storage duration with the `__cxa_atexit` function rather than the `atexit` function. This option is required for fully standards-compliant handling of static destructors, but will only work if your C library supports `__cxa_atexit`.

`-fvtable-thunks`

Use '`thunks`' to implement the virtual function dispatch table ('`vtable`'). The traditional (cfront-style) approach to implementing vtables was to store a pointer to the function and two offsets for adjusting the '`this`' pointer at the call site. Newer implementations store a single pointer to a '`thunk`' function which does any necessary adjustment and then calls the target function.

This option also enables a heuristic for controlling emission of vtables; if a class has any non-inline virtual functions, the vtable will be emitted in the translation unit containing the first one of those.

In the MIPS SDE compiler '`thunks`' are used by default. Like all options that change the ABI, all C++ code, *including libgcc.a* must be built with the same setting of this option.

`-nostdinc++`

Do not search for header files in the standard directories specific to C++, but do still search the other standard directories. (This option is used when building the C++ library.)

In addition, these optimization, warning, and code generation options have meanings only for C++ programs:

`-fno-default-inline`

Do not assume '`inline`' for functions defined inside a class scope. See Section 2.8 [Options That Control Optimization], page 33. Note that these functions will have linkage like inline functions; they just won't be inlined by default.

`-Wctor-dtor-privacy (C++ only)`

Warn when a class seems unusable, because all the constructors or destructors in a class are private and the class has no friends or public static member functions.

`-Wnon-virtual-dtor (C++ only)`

Warn when a class declares a non-virtual destructor that should probably be virtual, because it looks like the class will be used polymorphically.

`-Wreorder (C++ only)`

Warn when the order of member initializers given in the code does not match the order in which they must be executed. For instance:

```
struct A {
  int i;
```

```
  int j;
  A(): j (0), i (1) { }
};
```

Here the compiler will warn that the member initializers for ‘`i`’ and ‘`j`’ will be rearranged to match the declaration order of the members.

The following ‘`-W...`’ options are not affected by ‘`-Wall`’.

`-Weffc++` `(C++ only)`
: Warn about violations of various style guidelines from Scott Meyers’ *Effective C++* books. If you use this option, you should be aware that the standard library headers do not obey all of these guidelines; you can use ‘`grep -v`’ to filter out those warnings.

`-Wno-deprecated` `(C++ only)`
: Do not warn about usage of deprecated features. See Section 3.42 [Deprecated Features], page 113.

`-Wno-non-template-friend` `(C++ only)`
: Disable warnings when non-templatized friend functions are declared within a template. With the advent of explicit template specification support in g++, if the name of the friend is an unqualified-id (ie, ‘`friend foo(int)`’), the C++ language specification demands that the friend declare or define an ordinary, nontemplate function. (Section 14.5.3). Before g++ implemented explicit specification, unqualified-ids could be interpreted as a particular specialization of a templatized function. Because this non-conforming behavior is no longer the default behavior for g++, ‘`-Wnon-template-friend`’ allows the compiler to check existing code for potential trouble spots, and is on by default. This new compiler behavior can also be turned off with the flag ‘`-fguiding-decls`’, which activates the older, non-specification compiler code, or with ‘`-Wno-non-template-friend`’ which keeps the conformant compiler code but disables the helpful warning.

`-Wold-style-cast` `(C++ only)`
: Warn if an old-style (C-style) cast is used within a C++ program. The new-style casts (‘`static_cast`’, ‘`reinterpret_cast`’, and ‘`const_cast`’) are less vulnerable to unintended effects.

`-Woverloaded-virtual` `(C++ only)`
: Warn when a derived class function declaration may be an error in defining a virtual function. In a derived class, the definitions of virtual functions must match the type signature of a virtual function declared in the base class. With this option, the compiler warns when you define a function with the same name as a virtual function, but with a type signature that does not match any declarations from the base class.

`-Wno-pmf-conversions` `(C++ only)`
: Disable the diagnostic for converting a bound pointer to member function to a plain pointer.

`-Wsign-promo` `(C++ only)`
: Warn when overload resolution chooses a promotion from unsigned or enumeral type to a signed type over a conversion to an unsigned type of the same size. Previous versions of g++ would try to preserve unsignedness, but the standard mandates the current behavior.

`-Wsynth (C++ only)`

Warn when g++'s synthesis behavior does not match that of cfront. For instance:

```
struct A {
  operator int ();
  A& operator = (int);
};

main ()
{
  A a,b;
  a = b;
}
```

In this example, g++ will synthesize a default '`A& operator = (const A&);`', while cfront will use the user-defined '`operator =`'.

2.6 Options to Request or Suppress Warnings

Warnings are diagnostic messages that report constructions which are not inherently erroneous but which are risky or suggest there may have been an error.

You can request many specific warnings with options beginning '`-W`', for example '`-Wimplicit`' to request warnings on implicit declarations. Each of these specific warning options also has a negative form beginning '`-Wno-`' to turn off warnings; for example, '`-Wno-implicit`'. This manual lists only one of the two forms, whichever is not the default.

These options control the amount and kinds of warnings produced by GCC:

`-fsyntax-only`

Check the code for syntax errors, but don't do anything beyond that.

`-pedantic` Issue all the warnings demanded by strict ANSI C and ISO C++; reject all programs that use forbidden extensions.

Valid ANSI C and ISO C++ programs should compile properly with or without this option (though a rare few will require '`-ansi`'). However, without this option, certain GNU extensions and traditional C and C++ features are supported as well. With this option, they are rejected.

'`-pedantic`' does not cause warning messages for use of the alternate keywords whose names begin and end with '__'. Pedantic warnings are also disabled in the expression that follows `__extension__`. However, only system header files should use these escape routes; application programs should avoid them. See Section 3.36 [Alternate Keywords], page 109.

This option is not intended to be *useful*; it exists only to satisfy pedants who would otherwise claim that GCC fails to support the ANSI standard.

Some users try to use '`-pedantic`' to check programs for strict ANSI C conformance. They soon find that it does not do quite what they want: it finds some non-ANSI practices, but not all—only those for which ANSI C *requires* a diagnostic.

A feature to report any failure to conform to ANSI C might be useful in some instances, but would require considerable additional work and would be quite different from '`-pedantic`'. We don't have plans to support such a feature in the near future.

`-pedantic-errors`
: Like '`-pedantic`', except that errors are produced rather than warnings.

`-w`
: Inhibit all warning messages.

`-Wno-import`
: Inhibit warning messages about the use of '`#import`'.

`-Wchar-subscripts`
: Warn if an array subscript has type `char`. This is a common cause of error, as programmers often forget that this type is signed on some machines.

`-Wcomment`
: Warn whenever a comment-start sequence '`/*`' appears in a '`/*`' comment, or whenever a Backslash-Newline appears in a '`//`' comment.

`-Wformat`
: Check calls to `printf` and `scanf`, etc., to make sure that the arguments supplied have types appropriate to the format string specified.

`-Wimplicit-int`
: Warn when a declaration does not specify a type.

`-Wimplicit-function-declaration`
`-Werror-implicit-function-declaration`
: Give a warning (or error) whenever a function is used before being declared.

`-Wimplicit`
: Same as '`-Wimplicit-int`' and '`-Wimplicit-function-`' '`declaration`'.

`-Wmain`
: Warn if the type of '`main`' is suspicious. '`main`' should be a function with external linkage, returning int, taking either zero arguments, two, or three arguments of appropriate types.

`-Wmultichar`
: Warn if a multicharacter constant ('`'FOOF'`') is used. Usually they indicate a typo in the user's code, as they have implementation-defined values, and should not be used in portable code.

`-Wparentheses`
: Warn if parentheses are omitted in certain contexts, such as when there is an assignment in a context where a truth value is expected, or when operators are nested whose precedence people often get confused about.

  This option includes '`-Wparentheses-else`' below.

`-Wparentheses-else`
: Also warn about constructions where there may be confusion to which `if` statement an `else` branch belongs. Here is an example of such a case:

```
{
  if (a)
    if (b)
      foo ();
  else
    bar ();
```

```
}
```

In C, every `else` branch belongs to the innermost possible `if` statement, which in this example is `if (b)`. This is often not what the programmer expected, as illustrated in the above example by indentation the programmer chose. When there is the potential for this confusion, GNU C will issue a warning when this flag is specified. To eliminate the warning, add explicit braces around the innermost `if` statement so there is no way the `else` could belong to the enclosing `if`. The resulting code would look like this:

```
{
  if (a)
    {
      if (b)
        foo ();
      else
        bar ();
    }
}
```

`-Wreturn-type`
: Warn whenever a function is defined with a return-type that defaults to `int`. Also warn about any `return` statement with no return-value in a function whose return-type is not `void`.

`-Wswitch` Warn whenever a `switch` statement has an index of enumeral type and lacks a `case` for one or more of the named codes of that enumeration. (The presence of a `default` label prevents this warning.) `case` labels outside the enumeration range also provoke warnings when this option is used.

`-Wtrigraphs`
: Warn if any trigraphs are encountered (assuming they are enabled).

`-Wunused-function`
: Warn whenever a static function is declared but not defined or a non\-inline static function is unused.

`-Wunused-label`
: Warn whenever a label is declared but not used.

  To suppress this warning use the ‘`unused`’ attribute (see Section 3.29 [Variable Attributes], page 92).

`-Wunused-parameter`
: Warn whenever a function parameter is unused aside from its declaration.

  To suppress this warning use the ‘`unused`’ attribute (see Section 3.29 [Variable Attributes], page 92).

`-Wunused-variable`
: Warn whenever a local variable or non-constant static variable is unused aside from its declaration

  To suppress this warning use the ‘`unused`’ attribute (see Section 3.29 [Variable Attributes], page 92).

`-Wunused-value`

Warn whenever a statement computes a result that is explicitly not used.

To suppress this warning cast the expression to ‘`void`’.

`-Wunused` All all the above ‘`-Wunused`’ options combined.

In order to get a warning about an unused function parameter, you must either specify ‘`-W -Wunused`’ or separatly specify ‘`-Wunused-parameter`’.

`-Wuninitialized`

Warn if an automatic variable is used without first being initialized or if a variable may be clobbered by a `setjmp` call.

These warnings are possible only in optimizing compilation, because they require data flow information that is computed only when optimizing. If you don’t specify ‘`-O`’, you simply won’t get these warnings.

These warnings occur only for variables that are candidates for register allocation. Therefore, they do not occur for a variable that is declared `volatile`, or whose address is taken, or whose size is other than 1, 2, 4 or 8 bytes. Also, they do not occur for structures, unions or arrays, even when they are in registers.

Note that there may be no warning about a variable that is used only to compute a value that itself is never used, because such computations may be deleted by data flow analysis before the warnings are printed.

These warnings are made optional because GCC is not smart enough to see all the reasons why the code might be correct despite appearing to have an error. Here is one example of how this can happen:

```
{
  int x;
  switch (y)
    {
    case 1: x = 1;
      break;
    case 2: x = 4;
      break;
    case 3: x = 5;
    }
  foo (x);
}
```

If the value of `y` is always 1, 2 or 3, then `x` is always initialized, but GCC doesn’t know this. Here is another common case:

```
{
  int save_y;
  if (change_y) save_y = y, y = new_y;
  ...
  if (change_y) y = save_y;
}
```

This has no bug because `save_y` is used only if it is set.

This option also includes ‘`-Wlongjmp-clobbers`’ below.

`-Wlongjmp-clobbers`

Warn when a nonvolatile automatic variable might be changed by a call to `longjmp`. These warnings as well are possible only in optimizing compilation.

This option also warns when a nonvolatile automatic variable might be changed by a call to `longjmp`. These warnings as well are possible only in optimizing compilation.

The compiler sees only the calls to `setjmp`. It cannot know where `longjmp` will be called; in fact, a signal handler could call it at any point in the code. As a result, you may get a warning even when there is in fact no problem because `longjmp` cannot in fact be called at the place which would cause a problem.

Some spurious warnings can be avoided if you declare all the functions you use that never return as `noreturn`. See Section 3.23 [Function Attributes], page 86.

`-Wreorder (C++ only)`

Warn when the order of member initializers given in the code does not match the order in which they must be executed. For instance:

`-Wunknown-pragmas`

Warn when a #pragma directive is encountered which is not understood by GCC. If this command line option is used, warnings will even be issued for unknown pragmas in system header files. This is not the case if the warnings were only enabled by the '`-Wall`' command line option.

`-Wall` All of the above '`-W`' options combined. This enables all the warnings about constructions that some users consider questionable, and that are easy to avoid (or modify to prevent the warning), even in conjunction with macros.

The following '`-W...`' options are not implied by '`-Wall`'. Some of them warn about constructions that users generally do not consider questionable, but which occasionally you might wish to check for; others warn about constructions that are necessary or hard to avoid in some cases, and there is no simple way to modify the code to suppress the warning.

`-W` Print extra warning messages for these events:

- A function can return either with or without a value. (Falling off the end of the function body is considered returning without a value.) For example, this function would evoke such a warning:

```
foo (a)
{
  if (a > 0)
    return a;
}
```

- An expression-statement or the left-hand side of a comma expression contains no side effects. To suppress the warning, cast the unused expression to void. For example, an expression such as '`x[i,j]`' will cause a warning, but '`x[(void)i,j]`' will not.
- An unsigned value is compared against zero with '`<`' or '`<=`'.
- A comparison like '`x<=y<=z`' appears; this is equivalent to '`(x<=y ? 1 : 0) <= z`', which is a different interpretation from that of ordinary mathematical notation.
- Storage-class specifiers like `static` are not the first things in a declaration. According to the C Standard, this usage is obsolescent.

- If '`-Wall`' or '`-Wunused`' is also specified, warn about unused arguments.
- A comparison between signed and unsigned values could produce an incorrect result when the signed value is converted to unsigned. (But don't warn if '`-Wno-sign-compare`' is also specified.)
- An aggregate has a partly bracketed initializer. For example, the following code would evoke such a warning, because braces are missing around the initializer for `x.h`:

```
struct s { int f, g; };
struct t { struct s h; int i; };
struct t x = { 1, 2, 3 };
```

- An aggregate has an initializer which does not initialize all members. For example, the following code would cause such a warning, because `x.h` would be implicitly initialized to zero:

```
struct s { int f, g, h; };
struct s x = { 3, 4 };
```

`-Wfloat-equal`

Warn if floating point values are used in equality comparisons.

The idea behind this is that sometimes it is convenient (for the programmer) to consider floating-point values as approximations to infinitely precise real numbers. If you are doing this, then you need to compute (by analysing the code, or in some other way) the maximum or likely maximum error that the computation introduces, and allow for it when performing comparisons (and when producing output, but that's a different problem). In particular, instead of testing for equality, you would check to see whether the two values have ranges that overlap; and this is done with the relational operators, so equality comparisons are probably mistaken.

`-Wtraditional (C only)`

Warn about certain constructs that behave differently in traditional and ANSI C.

- Macro arguments occurring within string constants in the macro body. These would substitute the argument in traditional C, but are part of the constant in ANSI C.
- A function declared external in one block and then used after the end of the block.
- A `switch` statement has an operand of type `long`.
- A non-`static` function declaration follows a `static` one. This construct is not accepted by some traditional C compilers.
- The ANSI type of an integer constant has a different width or signedness from its traditional type. This warning is only issued if the base of the constant is ten. I.e. hexadecimal or octal values, which typically represent bit patterns, are not warned about.
- Usage of ANSI string concatenation is detected.

`-Wundef` Warn if an undefined identifier is evaluated in an '`#if`' directive.

`-Wshadow` Warn whenever a local variable shadows another local variable.

`-Wid-clash-`*len*
: Warn whenever two distinct identifiers match in the first *len* characters. This may help you prepare a program that will compile with certain obsolete, brain-damaged compilers.

`-Wlarger-than-`*len*
: Warn whenever an object of larger than *len* bytes is defined.

`-Wpointer-arith`
: Warn about anything that depends on the "size of" a function type or of `void`. GNU C assigns these types a size of 1, for convenience in calculations with `void *` pointers and pointers to functions.

`-Wbad-function-cast` `(C only)`
: Warn whenever a function call is cast to a non-matching type. For example, warn if `int malloc()` is cast to `anything *`.

`-Wcast-qual`
: Warn whenever a pointer is cast so as to remove a type qualifier from the target type. For example, warn if a `const char *` is cast to an ordinary `char *`.

`-Wcast-align`
: Warn whenever a pointer is cast such that the required alignment of the target is increased. For example, warn if a `char *` is cast to an `int *` on machines where integers can only be accessed at two- or four-byte boundaries.

`-Wwrite-strings`
: Give string constants the type `const char[`*length*`]` so that copying the address of one into a non-`const char *` pointer will get a warning. These warnings will help you find at compile time code that can try to write into a string constant, but only if you have been very careful about using `const` in declarations and prototypes. Otherwise, it will just be a nuisance; this is why we did not make '`-Wall`' request these warnings.

`-Wconversion`
: Warn if a prototype causes a type conversion that is different from what would happen to the same argument in the absence of a prototype. This includes conversions of fixed point to floating and vice versa, and conversions changing the width or signedness of a fixed point argument except when the same as the default promotion.

  Also, warn if a negative integer constant expression is implicitly converted to an unsigned type. For example, warn about the assignment `x = -1` if `x` is unsigned. But do not warn about explicit casts like `(unsigned) -1`.

`-Wsign-compare`
: Warn when a comparison between signed and unsigned values could produce an incorrect result when the signed value is converted to unsigned. This warning is also enabled by '`-W`'; to get the other warnings of '`-W`' without this warning, use '`-W -Wno-sign-compare`'.

`-Waggregate-return`
: Warn if any functions that return structures or unions are defined or called. (In languages where you can return an array, this also elicits a warning.)

`-Wstrict-prototypes` `(C only)`
Warn if a function is declared or defined without specifying the argument types. (An old-style function definition is permitted without a warning if preceded by a declaration which specifies the argument types.)

`-Wmissing-prototypes` `(C only)`
Warn if a global function is defined without a previous prototype declaration. This warning is issued even if the definition itself provides a prototype. The aim is to detect global functions that fail to be declared in header files.

`-Wmissing-declarations`
Warn if a global function is defined without a previous declaration. Do so even if the definition itself provides a prototype. Use this option to detect global functions that are not declared in header files.

`-Wmissing-noreturn`
Warn about functions which might be candidates for attribute `noreturn`. Note these are only possible candidates, not absolute ones. Care should be taken to manually verify functions actually do not ever return before adding the `noreturn` attribute, otherwise subtle code generation bugs could be introduced.

`-Wpacked` Warn if a structure is given the packed attribute, but the packed attribute has no effect on the layout or size of the structure. Such structures may be mis-aligned for little benefit. For instance, in this code, the variable `f.x` in `struct bar` will be misaligned even though `struct bar` does not itself have the packed attribute:

```
struct foo {
  int x;
  char a, b, c, d;
} __attribute__((packed));
struct bar {
  char z;
  struct foo f;
};
```

`-Wpadded` Warn if padding is included in a structure, either to align an element of the structure or to align the whole structure. Sometimes when this happens it is possible to rearrange the fields of the structure to reduce the padding and so make the structure smaller.

`-Wredundant-decls`
Warn if anything is declared more than once in the same scope, even in cases where multiple declaration is valid and changes nothing.

`-Wnested-externs` `(C only)`
Warn if an `extern` declaration is encountered within a function.

`-Wunreachable-code`
Warn if the compiler detects that code will never be executed.

This option is intended to warn when the compiler detects that at least a whole line of source code will never be executed, because some condition is never satisfied or because it is after a procedure that never returns.

It is possible for this option to produce a warning even though there are circumstances under which part of the affected line can be executed, so care should be taken when removing apparently-unreachable code.

For instance, when a function is inlined, a warning may mean that the line is unreachable in only one inlined copy of the function.

This option is not made part of '`-Wall`' because in a debugging version of a program there is often substantial code which checks correct functioning of the program and is, hopefully, unreachable because the program does work. Another common use of unreachable code is to provide behaviour which is selectable at compile-time.

`-Winline` Warn if a function can not be inlined and it was declared as inline.

`-Wlong-long`

Warn if '`long long`' type is used. This is default. To inhibit the warning messages, use '`-Wno-long-long`'. Flags '`-Wlong-long`' and '`-Wno-long-long`' are taken into account only when '`-pedantic`' flag is used.

`-Werror` Make all warnings into errors.

2.7 Options for Debugging Your Program or GCC

GCC has various special options that are used for debugging either your program or GCC:

`-g` Produce debugging information in the operating system's native format (stabs by default, ot DWARF if you ask for it). GDB can work with this debugging information.

On most systems that use stabs format, '`-g`' enables use of extra debugging information that only GDB can use; this extra information makes debugging work better in GDB but will probably make other debuggers crash or refuse to read the program. If you want to control for certain whether to generate the extra information, use '`-gstabs+`', '`-gstabs`', '`-gdwarf-1+`', or '`-gdwarf-1`' (see below).

Unlike most other C compilers, GCC allows you to use '`-g`' with '`-O`'. The shortcuts taken by optimized code may occasionally produce surprising results: some variables you declared may not exist at all; flow of control may briefly move where you did not expect it; some statements may not be executed because they compute constant results or their values were already at hand; some statements may execute in different places because they were moved out of loops.

Nevertheless it proves possible to debug optimized output. This makes it reasonable to use the optimizer for programs that might have bugs.

The following options are useful when GCC is generated with the capability for more than one debugging format.

`-ggdb` Produce debugging information for use by GDB. This means to use the most expressive format available (DWARF 2, stabs, or the native format if neither of those are supported), including GDB extensions if at all possible.

`-gstabs` Produce debugging information in stabs format (if that is supported), without GDB extensions. This is the format used by DBX on most BSD systems.

`-gstabs+` Produce debugging information in stabs format (if that is supported), using GNU extensions understood only by the GNU debugger (GDB). The use of these extensions is likely to make other debuggers crash or refuse to read the program.

`-gdwarf` Produce debugging information in DWARF version 1 format (if that is supported). This is the format used by SDB on most System V Release 4 systems.

`-gdwarf+` Produce debugging information in DWARF version 1 format (if that is supported), using GNU extensions understood only by the GNU debugger (GDB). The use of these extensions is likely to make other debuggers crash or refuse to read the program.

`-gdwarf-2` Produce debugging information in DWARF version 2 format (if that is supported). This is the format used by DBX on IRIX 6.

`-g`*level*
`-ggdb`*level*
`-gstabs`*level*
`-gdwarf`*level*
`-gdwarf-2`*level*

Request debugging information and also use *level* to specify how much information. The default level is 2.

Level 1 produces minimal information, enough for making backtraces in parts of the program that you don't plan to debug. This includes descriptions of functions and external variables, but no information about local variables and no line numbers.

Level 3 includes extra information, such as all the macro definitions present in the program. Some debuggers support macro expansion when you use '`-g3`'.

`-p` Generate extra code to write profile information suitable for the analysis program `prof`. You must use this option when compiling the source files you want data about, and you must also use it when linking.

`-pg` Generate extra code to write profile information suitable for the analysis program `gprof`. You must use this option when compiling the source files you want data about, and you must also use it when linking.

`-Q` Makes the compiler print out each function name as it is compiled, and print some statistics about each pass when it finishes.

`-fprofile-arcs`

Instrument *arcs* during compilation. For each function of your program, GCC creates a program flow graph, then finds a spanning tree for the graph. Only arcs that are not on the spanning tree have to be instrumented: the compiler adds code to count the number of times that these arcs are executed. When an arc is the only exit or only entrance to a block, the instrumentation code can be added to the block; otherwise, a new basic block must be created to hold the instrumentation code.

Since not every arc in the program must be instrumented, programs compiled with this option run faster than programs compiled with '`-a`', which adds instrumentation code to every basic block in the program. The tradeoff: since `gcov` does not have execution counts for all branches, it must start with the execution counts for the instrumented branches, and then iterate over the program flow graph until the entire graph has been solved. Hence, `gcov` runs a little more slowly than a program which uses information from '`-a`'.

'`-fprofile-arcs`' also makes it possible to estimate branch probabilities, and to calculate basic block execution counts. In general, basic block execution counts do not

give enough information to estimate all branch probabilities. When the compiled program exits, it saves the arc execution counts to a file called '*sourcename*`.da`'. Use the compiler option '`-fbranch-probabilities`' (see Section 2.8 [Options that Control Optimization], page 33) when recompiling, to optimize using estimated branch probabilities.

`-ftest-coverage`
Create data files for the `gcov` code-coverage utility (see Chapter 5 [`gcov`: a GCC Test Coverage Program], page 125). The data file names begin with the name of your source file:

sourcename`.bb`
A mapping from basic blocks to line numbers, which `gcov` uses to associate basic block execution counts with line numbers.

sourcename`.bbg`
A list of all arcs in the program flow graph. This allows `gcov` to reconstruct the program flow graph, so that it can compute all basic block and arc execution counts from the information in the *sourcename*`.da` file (this last file is the output from '`-fprofile-arcs`').

`-d`*letters* Says to make debugging dumps during compilation at times specified by *letters*. This is used for debugging the compiler. The file names for most of the dumps are made by appending a pass number and a word to the source file name (e.g. '`foo.c.00.rtl`' or '`foo.c.01.jump`'). Here are the possible letters for use in *letters*, and their meanings:

'`A`' Annotate the assembler output with miscellaneous debugging information.

'`b`' Dump after computing branch probabilities, to '*file*`.10.bp`'.

'`B`' Dump after block reordering, to '*file*`.20.bbro`'.

'`c`' Dump after instruction combination, to the file '*file*`.12.combine`'.

'`d`' Dump after delayed branch scheduling, to '*file*`.24.dbr`'.

'`D`' Dump all macro definitions, at the end of preprocessing, in addition to normal output.

'`e`' Dump after SSA optimizations, to '*file*`.05.ssa`' and '*file*`.06.ussa`'.

'`f`' Dump after flow analysis, to '*file*`.11.flow`'.

'`F`' Dump after purging `ADDRESSOF` codes, to '*file*`.04.addressof`'.

'`g`' Dump after global register allocation, to '*file*`.16.greg`'.

'`G`' Dump after GCSE, to '*file*`.07.gcse`'.

'`i`' Dump after sibling call optimizations, to '*file*`.01.sibling`'.

'`j`' Dump after first jump optimization, to '*file*`.02.jump`'.

'`J`' Dump after last jump optimization, to '*file*`.22.jump2`'.

'`k`' Dump after conversion from registers to stack, to '*file*`.25.stack`'.

‘`l`’ Dump after local register allocation, to ‘*file*`.15.lreg`’.

‘`L`’ Dump after loop optimization, to ‘*file*`.08.loop`’.

‘`M`’ Dump after performing the machine dependent reorganisation pass, to ‘*file*`.23.mach`’.

‘`n`’ Dump after register renumbering, to ‘*file*`.21.rnreg`’.

‘`N`’ Dump after the register move pass, to ‘*file*`.13.regmove`’.

‘`r`’ Dump after RTL generation, to ‘*file*`.00.rtl`’.

‘`R`’ Dump after the second instruction scheduling pass, to ‘*file*`.19.sched2`’.

‘`s`’ Dump after CSE (including the jump optimization that sometimes follows CSE), to ‘*file*`.03.cse`’.

‘`S`’ Dump after the first instruction scheduling pass, to ‘*file*`.14.sched`’.

‘`t`’ Dump after the second CSE pass (including the jump optimization that sometimes follows CSE), to ‘*file*`.09.cse2`’.

‘`w`’ Dump after the second flow pass, to ‘*file*`.17.flow2`’.

‘`z`’ Dump after the peephole pass, to ‘*file*`.18.peephole2`’.

‘`a`’ Produce all the dumps listed above.

‘`m`’ Print statistics on memory usage, at the end of the run, to standard error.

‘`p`’ Annotate the assembler output with a comment indicating which pattern and alternative was used. The length of each instruction is also printed.

‘`v`’ For each of the other indicated dump files (except for ‘*file*`.00.rtl`’), dump a representation of the control flow graph suitible for viewing with VCG to ‘*file*`.`*pass*`.vcg`’.

‘`w`’ Dump after the second flow pass to ‘*file*`.14.flow2`’.

‘`x`’ Just generate RTL for a function instead of compiling it. Usually used with ‘`r`’.

‘`y`’ Dump debugging information during parsing, to standard error.

‘`z`’ Dump after the peephole2 pass to ‘*file*`.15.peephole2`’.

`-fdump-unnumbered`

When doing debugging dumps (see -d option above), suppress instruction numbers and line number note output. This makes it more feasible to use diff on debugging dumps for compiler invokations with different options, in particular with and without -g.

`-fdump-translation-unit-`*file* `(C++ only)`

Dump a representation of the tree structure for the entire translation unit to *file*.

`-fpretend-float`

When running a cross-compiler, pretend that the target machine uses the same floating point format as the host machine. This causes incorrect output of the actual floating constants, but the actual instruction sequence will probably be the same as GCC would make when running on the target machine.

`-save-temps`
: Store the usual "temporary" intermediate files permanently; place them in the current directory and name them based on the source file. Thus, compiling '`foo.c`' with '`-c -save-temps`' would produce files '`foo.i`' and '`foo.s`', as well as '`foo.o`'.

`-time`
: Report the CPU time taken by each subprocess in the compilation sequence. For C source files, this is the preprocessor, compiler proper, and assembler. The output looks like this:

```
# cpp 0.04 0.04
# cc1 0.12 0.01
# as 0.00 0.01
```

The first number on each line is the "user time," that is time spent executing the program itself. The second number is "system time," time spent executing operating system routines on behalf of the program. Both numbers are in seconds.

`-print-file-name=`*library*
: Print the full absolute name of the library file *library* that would be used when linking—and don't do anything else. With this option, GCC does not compile or link anything; it just prints the file name.

`-print-prog-name=`*program*
: Like '`-print-file-name`', but searches for a program such as '`cpp`'.

`-print-libgcc-file-name`
: Same as '`-print-file-name=libgcc.a`'.

This is useful when you use '`-nostdlib`' or '`-nodefaultlibs`' but you do want to link with '`libgcc.a`'. You can do

```
gcc -nostdlib files... `gcc -print-libgcc-file-name`
```

`-print-search-dirs`
: Print the name of the configured installation directory and a list of program and library directories gcc will search—and don't do anything else.

This is useful when gcc prints the error message '`installation problem, cannot exec cpp: No such file or directory`'. To resolve this you either need to put '`cpp`' and the other compiler components where gcc expects to find them, or you can set the environment variable `GCC_EXEC_PREFIX` to the directory where you installed them. Don't forget the trailing '/'. See Section 2.17 [Environment Variables], page 67.

2.8 Options That Control Optimization

These options control various sorts of optimizations:

`-O`
`-O1`
: Optimize. Optimizing compilation takes somewhat more time, and a lot more memory for a large function.

Without '`-O`', the compiler's goal is to reduce the cost of compilation and to make debugging produce the expected results. Statements are independent: if you stop the program with a breakpoint between statements, you can then assign a new value to any

variable or change the program counter to any other statement in the function and get exactly the results you would expect from the source code.

Without ‘`-O`’, the compiler only allocates variables declared `register` in registers. The resulting compiled code is a little worse than produced by PCC without ‘`-O`’.

With ‘`-O`’, the compiler tries to reduce code size and execution time.

When you specify ‘`-O`’, the compiler turns on ‘`-fthread-jumps`’ and ‘`-fdefer-pop`’ on all machines. The compiler turns on ‘`-fdelayed-branch`’ on machines that have delay slots, and ‘`-fomit-frame-pointer`’ on machines that can support debugging even without a frame pointer. On some machines the compiler also turns on other flags.

`-O2` Optimize even more. GCC performs nearly all supported optimizations that do not involve a space-speed tradeoff. The compiler does not perform loop unrolling or function inlining when you specify ‘`-O2`’. As compared to ‘`-O`’, this option increases both compilation time and the performance of the generated code.

‘`-O2`’ turns on all optional optimizations except for loop unrolling and function inlining. It also turns on the ‘`-fforce-mem`’ option on all machines and frame pointer elimination on machines where doing so does not interfere with debugging.

`-O3` Optimize yet more. ‘`-O3`’ turns on all optimizations specified by ‘`-O2`’ and also turns on the ‘`inline-functions`’ option.

`-O0` Do not optimize.

`-Os` Optimize for size. ‘`-Os`’ enables all ‘`-O2`’ optimizations that do not typically increase code size. It also performs further optimizations designed to reduce code size.

If you use multiple ‘`-O`’ options, with or without level numbers, the last such option is the one that is effective.

Options of the form ‘`-f`*flag*’ specify machine-independent flags. Most flags have both positive and negative forms; the negative form of ‘`-ffoo`’ would be ‘`-fno-foo`’. In the table below, only one of the forms is listed—the one which is not the default. You can figure out the other form by either removing ‘`no-`’ or adding it.

`-ffloat-store`

Do not store floating point variables in registers, and inhibit other options that might change whether a floating point value is taken from a register or memory.

This option prevents undesirable excess precision on machines such as the 68000 where the floating registers (of the 68881) keep more precision than a `double` is supposed to have. Similarly for the x86 architecture. For most programs, the excess precision does only good, but a few programs rely on the precise definition of IEEE floating point. Use ‘`-ffloat-store`’ for such programs, after modifying them to store all pertinent intermediate computations into variables.

`-fno-default-inline`

Do not make member functions inline by default merely because they are defined inside the class scope (C++ only). Otherwise, when you specify ‘`-O`’, member functions defined inside class scope are compiled inline by default; i.e., you don’t need to add ‘`inline`’ in front of the member function name.

`-fno-defer-pop`

Always pop the arguments to each function call as soon as that function returns. For machines which must pop arguments after a function call, the compiler normally lets arguments accumulate on the stack for several function calls and pops them all at once.

`-fforce-mem`

Force memory operands to be copied into registers before doing arithmetic on them. This produces better code by making all memory references potential common subexpressions. When they are not common subexpressions, instruction combination should eliminate the separate register-load. The '`-O2`' option turns on this option.

`-fforce-addr`

Force memory address constants to be copied into registers before doing arithmetic on them. This may produce better code just as '`-fforce-mem`' may.

`-fomit-frame-pointer`

Don't keep the frame pointer in a register for functions that don't need one. An (optimised) MIPS function will only use a frame pointer when it does something funny with the stack, perhaps because it contains a call to '`alloca()`' or a variable-length array declaration.

On some machines, such as the Vax, this flag has no effect, because the standard calling sequence automatically handles the frame pointer and nothing is saved by pretending it doesn't exist. The machine-description macro `FRAME_POINTER_REQUIRED` controls whether a target machine supports this flag. See section "Register Usage" in *Using and Porting GCC*.

`-foptimize-sibling-calls`

Optimize sibling and tail recursive calls.

`-fno-inline`

Don't pay attention to the `inline` keyword. Normally this option is used to keep the compiler from expanding any functions inline. Note that if you are not optimizing, no functions can be expanded inline.

`-finline-functions`

Integrate all simple functions into their callers. The compiler heuristically decides which functions are simple enough to be worth integrating in this way.

If all calls to a given function are integrated, and the function is declared `static`, then the function is normally not output as assembler code in its own right.

`-finline-limit=`*n*

By default, gcc limits the size of functions that can be inlined. This flag allows the control of this limit for functions that are explicitly marked as inline (ie marked with the inline keyword or defined within the class definition in c++). *n* is the size of functions that can be inlined in number of pseudo instructions (not counting parameter handling). The default value of n is 10000. Increasing this value can result in more inlined code at the cost of compilation time and memory consumption. Decreasing usually makes the compilation faster and less code will be inlined (which presumably means slower programs). This option is particularly useful for programs that use inlining heavily such as those based on recursive templates with c++.

Note: pseudo instruction represents, in this particular context, an abstract measurement of function's size. In no way, it represents a count of assembly instructions and as such its exact meaning might change from one release to an another.

`-fkeep-inline-functions`
: Even if all calls to a given function are integrated, and the function is declared `static`, nevertheless output a separate run-time callable version of the function. This switch does not affect `extern inline` functions.

`-fkeep-static-consts`
: Emit variables declared `static const` when optimization isn't turned on, even if the variables aren't referenced.

  GCC enables this option by default. If you want to force the compiler to check if the variable was referenced, regardless of whether or not optimization is turned on, use the '`-fno-keep-static-consts`' option.

`-fno-function-cse`
: Do not put function addresses in registers; make each instruction that calls a constant function contain the function's address explicitly.

  This option results in less efficient code, but some strange hacks that alter the assembler output may be confused by the optimizations performed when this option is not used.

`-ffast-math`
: This option allows GCC to violate some ANSI or IEEE rules and/or specifications in the interest of optimizing code for speed. For example, it allows the compiler to assume arguments to the `sqrt` function are non-negative numbers and that no floating-point values are NaNs.

  Where the CPU provides fast but non-IEEE-compliant instructions (like MIPS IV's reciprocal and reciprocal square root) the compiler takes this flag as an OK to use them.

  This option should never be turned on by any '`-O`' option since it can result in incorrect output for programs which depend on an exact implementation of IEEE or ANSI rules/specifications for math functions.

`-fno-math-errno`
: Do not set ERRNO after calling math functions that are executed with a single instruction, e.g., sqrt. A program that relies on IEEE exceptions for math error handling may want to use this flag for speed while maintaining IEEE arithmetic compatibility.

  The default is '`-fmath-errno`'. The '`-ffast-math`' option sets '`-fno-math-errno`'.

The following options control specific optimizations. The '`-O2`' option turns on all of these optimizations except '`-funroll-loops`' and '`-funroll-all-loops`'. On most machines, the '`-O`' option turns on the '`-fthread-jumps`' and '`-fdelayed-branch`' options, but specific machines may handle it differently.

You can use the following flags in the rare cases when "fine-tuning" of optimizations to be performed is desired.

`-fstrength-reduce`
: Perform the optimizations of loop strength reduction and elimination of iteration variables.

`-fthread-jumps`
: Perform optimizations where we check to see if a jump branches to a location where another comparison subsumed by the first is found. If so, the first branch is redirected to either the destination of the second branch or a point immediately following it, depending on whether the condition is known to be true or false.

`-fcse-follow-jumps`
: In common subexpression elimination, scan through jump instructions when the target of the jump is not reached by any other path. For example, when CSE encounters an `if` statement with an `else` clause, CSE will follow the jump when the condition tested is false.

`-fcse-skip-blocks`
: This is similar to '`-fcse-follow-jumps`', but causes CSE to follow jumps which conditionally skip over blocks. When CSE encounters a simple `if` statement with no else clause, '`-fcse-skip-blocks`' causes CSE to follow the jump around the body of the `if`.

`-frerun-cse-after-loop`
: Re-run common subexpression elimination after loop optimizations has been performed.

`-frerun-loop-opt`
: Run the loop optimizer twice.

`-fgcse`
: Perform a global common subexpression elimination pass. This pass also performs global constant and copy propagation.

`-flive-range`
: Perform live range splitting of variables at loop boundaries. This option is enabled by default at '`-O2`' optimization and higher for targets which use stabs debug symbols.

`-fdelete-null-pointer-checks`
: Use global dataflow analysis to identify and eliminate useless null pointer checks. Programs which rely on NULL pointer dereferences *not* halting the program may not work properly with this option. Use -fno-delete-null-pointer-checks to disable this optimizing for programs which depend on that behavior.

`-fexpensive-optimizations`
: Perform a number of minor optimizations that are relatively expensive.

`-foptimize-register-moves`
`-fregmove`
: Attempt to reassign register numbers in move instructions and as operands of other simple instructions in order to maximize the amount of register tying. This is especially helpful on machines with two-operand instructions. GCC enables this optimization by default with '`-O2`' or higher.

  Note `-fregmove` and `-foptimize-register-moves` are the same optimization.

`-fdelayed-branch`
: If supported for the target machine, attempt to reorder instructions to exploit instruction slots available after delayed branch instructions.

`-fschedule-insns`
: If supported for the target machine, attempt to reorder instructions to eliminate execution stalls due to required data being unavailable. This helps machines that have slow floating point or memory load instructions by allowing other instructions to be issued until the result of the load or floating point instruction is required.

`-fschedule-insns2`
: Similar to '`-fschedule-insns`', but requests an additional pass of instruction scheduling after register allocation has been done. This is especially useful on machines with a relatively small number of registers and where memory load instructions take more than one cycle.

`-ffunction-sections`
`-fdata-sections`
: Place each function or data item into its own section in the output file if the target supports arbitrary sections. The name of the function or the name of the data item determines the section's name in the output file.

  Use these options on systems where the linker can perform optimizations to improve locality of reference in the instruction space. HPPA processors running HP-UX and Sparc processors running Solaris 2 have linkers with such optimizations. Other systems using the ELF object format as well as AIX may have these optimizations in the future.

  Only use these options when there are significant benefits from doing so. When you specify these options, the assembler and linker will create larger object and executable files and will also be slower. You will not be able to use `gprof` on all systems if you specify this option and you may have problems with debugging if you specify both this option and '`-g`'.

`-fcaller-saves`
: Enable values to be allocated in registers that will be clobbered by function calls, by emitting extra instructions to save and restore the registers around such calls. Such allocation is done only when it seems to result in better code than would otherwise be produced.

  This option is always enabled by default on certain machines, usually those which have no call-preserved registers to use instead.

  For all machines, optimization level 2 and higher enables this flag by default.

`-funroll-loops`
: Perform the optimization of loop unrolling. This is only done for loops whose number of iterations can be determined at compile time or run time. '`-funroll-loops`' implies both '`-fstrength-reduce`' and '`-frerun-cse-after-loop`'.

`-funroll-all-loops`
: Perform the optimization of loop unrolling. This is done for all loops and usually makes programs run more slowly. '`-funroll-all-loops`' implies '`-fstrength-reduce`' as well as '`-frerun-cse-after-loop`'.

`-fmove-all-movables`
: Forces all invariant computations in loops to be moved outside the loop.

`-freduce-all-givs`

Forces all general-induction variables in loops to be strength-reduced.

Note: When compiling programs written in Fortran, '`-fmove-all-movables`' and '`-freduce-all-givs`' are enabled by default when you use the optimizer.

These options may generate better or worse code; results are highly dependent on the structure of loops within the source code.

These two options are intended to be removed someday, once they have helped determine the efficacy of various approaches to improving loop optimizations.

Please let us (`support@mips.com` and/or `gcc@gcc.gnu.org`) know how use of these options affects the performance of your production code. We're very interested in code that runs *slower* when these options are *enabled*.

`-fno-peephole`

Disable any machine-specific peephole optimizations.

`-fbranch-probabilities`

After running a program compiled with '`-fprofile-arcs`' (see Section 2.7 [Options for Debugging Your Program or `gcc`], page 29), you can compile it a second time using '`-fbranch-probabilities`', to improve optimizations based on guessing the path a branch might take.

`-fstrict-aliasing`

Allows the compiler to assume the strictest aliasing rules applicable to the language being compiled. For C (and C++), this activates optimizations based on the type of expressions. In particular, an object of one type is assumed never to reside at the same address as an object of a different type, unless the types are almost the same. For example, an `unsigned int` can alias an `int`, but not a `void*` or a `double`. A character type may alias any other type.

Pay special attention to code like this:

```
union a_union {
  int i;
  double d;
};

int f() {
  a_union t;
  t.d = 3.0;
  return t.i;
}
```

The practice of reading from a different union member than the one most recently written to (called "type-punning") is common. Even with '`-fstrict-aliasing`', type-punning is allowed, provided the memory is accessed through the union type. So, the code above will work as expected. However, this code might not:

```
int f() {
  a_union t;
  int* ip;
  t.d = 3.0;
  ip = &t.i;
```

```
  return *ip;
}
```

`-falign-functions`
`-falign-functions=`*n*
: Align the start of functions to the next power-of-two greater than *n*, skipping up to *n* bytes. For instance, '`-falign-functions=32`' aligns functions to the next 32-byte boundary, but '`-falign-functions=24`' would align to the next 32-byte boundary only if this can be done by skipping 23 bytes or less.

  '`-fno-align-functions`' and '`-falign-functions=1`' are equivalent and mean that functions will not be aligned.

  Some assemblers only support this flag when *n* is a power of two; in that case, it is rounded up.

  If *n* is not specified, use a machine-dependent default.

`-falign-labels`
`-falign-labels=`*n*
: Align all branch targets to a power-of-two boundary, skipping up to *n* bytes like '`-falign-functions`'. This option can easily make code slower, because it must insert dummy operations for when the branch target is reached in the usual flow of the code.

  If '`-falign-loops`' or '`-falign-jumps`' are applicable and are greater than this value, then their values are used instead.

  If *n* is not specified, use a machine-dependent default which is very likely to be '`1`', meaning no alignment.

`-falign-loops`
`-falign-loops=`*n*
: Align loops to a power-of-two boundary, skipping up to *n* bytes like '`-falign-functions`'. The hope is that the loop will be executed many times, which will make up for any execution of the dummy operations.

  If *n* is not specified, use a machine-dependent default.

`-falign-jumps`
`-falign-jumps=`*n*
: Align branch targets to a power-of-two boundary, for branch targets where the targets can only be reached by jumping, skipping up to *n* bytes like '`-falign-functions`'. In this case, no dummy operations need be executed.

  If *n* is not specified, use a machine-dependent default.

`-fssa`
: Perform optimizations in static single assignment form. Each function's flow graph is translated into SSA form, optimizations are performed, and the flow graph is translated back from SSA form. (Currently, no SSA-based optimizations are implemented, but converting into and out of SSA form is not an invariant operation, and generated code may differ.)

2.9 Options Controlling the Preprocessor

These options control the C preprocessor, which is run on each C source file before actual compilation.

If you use the '`-E`' option, nothing is done except preprocessing. Some of these options make sense only together with '`-E`' because they cause the preprocessor output to be unsuitable for actual compilation.

`-include` *file*
: Process *file* as input before processing the regular input file. In effect, the contents of *file* are compiled first. Any '`-D`' and '`-U`' options on the command line are always processed before '`-include` *file*', regardless of the order in which they are written. All the '`-include`' and '`-imacros`' options are processed in the order in which they are written.

`-imacros` *file*
: Process *file* as input, discarding the resulting output, before processing the regular input file. Because the output generated from *file* is discarded, the only effect of '`-imacros` *file*' is to make the macros defined in *file* available for use in the main input.

  Any '`-D`' and '`-U`' options on the command line are always processed before '`-imacros` *file*', regardless of the order in which they are written. All the '`-include`' and '`-imacros`' options are processed in the order in which they are written.

`-idirafter` *dir*
: Add the directory *dir* to the second include path. The directories on the second include path are searched when a header file is not found in any of the directories in the main include path (the one that '`-I`' adds to).

`-iprefix` *prefix*
: Specify *prefix* as the prefix for subsequent '`-iwithprefix`' options.

`-iwithprefix` *dir*
: Add a directory to the second include path. The directory's name is made by concatenating *prefix* and *dir*, where *prefix* was specified previously with '`-iprefix`'. If you have not specified a prefix yet, the directory containing the installed passes of the compiler is used as the default.

`-iwithprefixbefore` *dir*
: Add a directory to the main include path. The directory's name is made by concatenating *prefix* and *dir*, as in the case of '`-iwithprefix`'.

`-isystem` *dir*
: Add a directory to the beginning of the second include path, marking it as a system directory, so that it gets the same special treatment as is applied to the standard system directories.

`-isystem-c++` *dir*
: Same behavior as with '`-isystem`', but do not make headers in *dir* be implicitly evaluated as if they include the '`extern "C"`' linkage specification.

`-nostdinc` Do not search the standard system directories for header files. Only the directories you have specified with '`-I`' options (and the current directory, if appropriate) are searched. See Section 2.12 [Directory Options], page 46, for information on '`-I`'.

  By using both '`-nostdinc`' and '`-I-`', you can limit the include-file search path to only those directories you specify explicitly.

`-undef` Do not predefine any nonstandard macros. (Including architecture flags).

`-E` Run only the C preprocessor. Preprocess all the C source files specified and output the results to standard output or to the specified output file.

`-C` Tell the preprocessor not to discard comments. Used with the '`-E`' option.

`-P` Tell the preprocessor not to generate '`#line`' directives. Used with the '`-E`' option.

`-M` Tell the preprocessor to output a rule suitable for `make` describing the dependencies of each object file. For each source file, the preprocessor outputs one `make`-rule whose target is the object file name for that source file and whose dependencies are all the `#include` header files it uses. This rule may be a single line or may be continued with '`\`'-newline if it is long. The list of rules is printed on standard output instead of the preprocessed C program.

'`-M`' implies '`-E`'.

Another way to specify output of a `make` rule is by setting the environment variable `DEPENDENCIES_OUTPUT` (see Section 2.17 [Environment Variables], page 67).

`-MM` Like '`-M`' but the output mentions only the user header files included with '`#include "`*file*`"`'. System header files included with '`#include <`*file*`>`' are omitted.

`-MD` Like '`-M`' but the dependency information is written to a file made by replacing `".c"` with `".d"` at the end of the input file names. This is in addition to compiling the file as specified—'`-MD`' does not inhibit ordinary compilation the way '`-M`' does.

In Mach, you can use the utility `md` to merge multiple dependency files into a single dependency file suitable for using with the '`make`' command.

`-MMD` Like '`-MD`' except mention only user header files, not system header files.

`-MG` Treat missing header files as generated files and assume they live in the same directory as the source file. If you specify '`-MG`', you must also specify either '`-M`' or '`-MM`'. '`-MG`' is not supported with '`-MD`' or '`-MMD`'.

`-H` Print the name of each header file used, in addition to other normal activities.

`-A`*question*`(`*answer*`)`
Assert the answer *answer* for *question*, in case it is tested with a preprocessing conditional such as '`#if #`*question*`(`*answer*`)`'. '`-A-`' disables the standard assertions that normally describe the target machine.

`-D`*macro* Define macro *macro* with the string '`1`' as its definition.

`-D`*macro*`=`*defn*
Define macro *macro* as *defn*. All instances of '`-D`' on the command line are processed before any '`-U`' options.

`-U`*macro* Undefine macro *macro*. '`-U`' options are evaluated after all '`-D`' options, but before any '`-include`' and '`-imacros`' options.

`-dM` Tell the preprocessor to output only a list of the macro definitions that are in effect at the end of preprocessing. Used with the '`-E`' option.

`-dD` Tell the preprocessing to pass all macro definitions into the output, in their proper sequence in the rest of the output.

`-dN` Like '`-dD`' except that the macro arguments and contents are omitted. Only '`#define` *name*' is included in the output.

`-trigraphs`
Support ANSI C trigraphs. The '`-ansi`' option also has this effect.

`-Wp,`*option* Pass *option* as an option to the preprocessor. If *option* contains commas, it is split into multiple options at the commas.

2.10 Passing Options to the Assembler

You can pass options to the assembler.

`-Wa,`*option* Pass *option* as an option to the assembler. If *option* contains commas, it is split into multiple options at the commas.

2.11 Options for Linking

These options come into play when the compiler links object files into an executable output file. They are meaningless if the compiler is not doing a link step.

object-file-name
A file name that does not end in a special recognized suffix is considered to name an object file or library. (Object files are distinguished from libraries by the linker according to the file contents.) If linking is done, these object files are used as input to the linker.

`-c`
`-S`
`-E` If any of these options is used, then the linker is not run, and object file names should not be used as arguments. See Section 2.2 [Overall Options], page 8.

`-l`*library* Search the library named *library* when linking.

It makes a difference where in the command you write this option; the linker searches processes libraries and object files in the order they are specified. Thus, '`foo.o -lz bar.o`' searches library '`z`' after file '`foo.o`' but before '`bar.o`'. If '`bar.o`' refers to functions in '`z`', those functions may not be loaded.

The linker searches a standard list of directories for the library, which is actually a file named '`lib`*library*`.a`'. The linker then uses this file as if it had been specified precisely by name.

The directories searched include several standard system directories plus any that you specify with '`-L`'.

But since this is a cross-compiler for multiple targets, there are some flags to the compiler which produce incompatible code; the actual library files you used depend on those flags. The "multi-lib" feature puts a different library root directory in the search path to match what your flags have asked for.

The flags which affect the choice of library are:

`-EB, -EL` The "endianness" the software is built for.

`-mips1, -mips2`
: Libraries built using the MIPS I instruction set.

`-mips3, -mips4`
: Libraries built using the MIPS III instruction set.

`-mips32` Libraries built using MIPS32 instruction set.

`-mips32r2` Libraries built using MIPS32 release 2 instruction set.

`-mips64, -mips64r2`
: Libraries built using MIPS64 instruction set.

`-mips16, -mips16e`
: Libraries built using the MIPS16 or MIPS16e instruction set extension, depending on the base ISA.

`-mgp32` When combined with '`-mips3`' or '`-mips4`' selects libraries built with 32-bit integer registers, but using the full set of 32 64-bit floating-point registers.

`-mfp64` When combined with '`-mips32r2`' only, selects libraries built with 32-bit integer instructions, but using the full set of 32 64-bit floating-point registers, and the MIPS64 extended floating point instructions.

`-mno-float`
: Floating-point-free libraries (smaller).

`-msoft-float`
: Built with all floating-point operations as subroutine calls to a software floating-point emulation library.

`-msingle-float`
: Built using only single-precision floating-point instructions, with double-precision implemented as subroutine calls (like '`-msoft-float`'); probably only useful for the IDT R4640 and R4650 CPUs.

You might reasonably expect there to be different libraries for '`-G0`' and '`-G8`' but there aren't; '`-G0`' might in theory be less efficient, but the effect is very small and its completely compatible, so all the libraries are built that way.

Normally the files found this way are library files—archive files whose members are object files. The linker handles an archive file by scanning through it for members which define symbols that have so far been referenced but not defined. But if the file that is found is an ordinary object file, it is linked in the usual fashion. The only difference between using an '`-l`' option and specifying a file name is that '`-l`' surrounds *library* with '`lib`' and '`.a`' and searches several directories.

`-lobjc` You need this special case of the '`-l`' option in order to link an Objective C program.

`-nostartfiles`
: Do not use the standard system startup files when linking. The standard system libraries are used normally, unless `-nostdlib` or `-nodefaultlibs` is used.

`-nodefaultlibs`

Do not use the standard system libraries when linking. Only the libraries you specify will be passed to the linker. The standard startup files are used normally, unless `-nostartfiles` is used. The compiler may generate calls to memcmp, memset, and memcpy for System V (and ANSI C) environments or to bcopy and bzero for BSD environments. These entries are usually resolved by entries in libc. These entry points should be supplied through some other mechanism when this option is specified.

`-nostdlib` Do not use the standard system startup files or libraries when linking. No startup files and only the libraries you specify will be passed to the linker. The compiler may generate calls to memcmp, memset, and memcpy for System V (and ANSI C) environments or to bcopy and bzero for BSD environments. These entries are usually resolved by entries in libc. These entry points should be supplied through some other mechanism when this option is specified.

One of the standard libraries bypassed by '`-nostdlib`' and '`-nodefaultlibs`' is '`libgcc.a`', a library of internal subroutines that GCC uses to overcome shortcomings of particular machines, or special needs for some languages. (See section "Interfacing to GCC Output" in *Porting GCC*, for more discussion of '`libgcc.a`'.) In most cases, you need '`libgcc.a`' even when you want to avoid other standard libraries. In other words, when you specify '`-nostdlib`' or '`-nodefaultlibs`' you should usually specify '`-lgcc`' as well. This ensures that you have no unresolved references to internal GCC library subroutines.

`-s` Remove all symbol table and relocation information from the executable.

`-static` On systems that support dynamic linking, this prevents linking with the shared libraries. On other systems, this option has no effect.

`-shared` Produce a shared object which can then be linked with other objects to form an executable. Not all systems support this option. You must also specify '`-fpic`' or '`-fPIC`' on some systems when you specify this option.

`-symbolic` Bind references to global symbols when building a shared object. Warn about any unresolved references (unless overridden by the link editor option '`-Xlinker -z -Xlinker defs`'). Only a few systems support this option.

`-Xlinker` *option*

Pass *option* as an option to the linker. You can use this to supply system-specific linker options which GCC does not know how to recognize.

If you want to pass an option that takes an argument, you must use '`-Xlinker`' twice, once for the option and once for the argument. For example, to pass '`-assert definitions`', you must write '`-Xlinker -assert -Xlinker definitions`'. It does not work to write '`-Xlinker "-assert definitions"`', because this passes the entire string as a single argument, which is not what the linker expects.

`-Wl,`*option* Pass *option* as an option to the linker. If *option* contains commas, it is split into multiple options at the commas.

`-u` *symbol* Pretend the symbol *symbol* is undefined, to force linking of library modules to define it. You can use '`-u`' multiple times with different symbols to force loading of additional library modules.

2.12 Options for Directory Search

These options specify directories to search for header files, for libraries and for parts of the compiler:

`-I`*dir* 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 header file, substituting your own version, since these directories are searched before the system header file directories. If you use more than one '`-I`' option, the directories are scanned in left-to-right order; the standard system directories come after.

`-I-` Any directories you specify with '`-I`' options before the '`-I-`' option are searched only for the case of '`#include "`*file*`"`'; they are not searched for '`#include <`*file*`>`'.

If additional directories are specified with '`-I`' options after the '`-I-`', these directories are searched for all '`#include`' directives. (Ordinarily *all* '`-I`' directories are used this way.)

In addition, the '`-I-`' option inhibits the use of the current directory (where the current input file came from) as the first search directory for '`#include "`*file*`"`'. There is no way to override this effect of '`-I-`'. With '`-I.`' you can specify searching the directory which was current when the compiler was invoked. That is not exactly the same as what the preprocessor does by default, but it is often satisfactory.

'`-I-`' does not inhibit the use of the standard system directories for header files. Thus, '`-I-`' and '`-nostdinc`' are independent.

`-L`*dir* Add directory *dir* to the list of directories to be searched for '`-l`'.

`-B`*prefix* This option specifies where to find the executables, libraries, include files, and data files of the compiler itself.

The compiler driver program runs one or more of the subprograms '`cpp`', '`cc1`', '`as`' and '`ld`'. It tries *prefix* as a prefix for each program it tries to run.

For each subprogram to be run, the compiler driver first tries the '`-B`' prefix, if any. If that name is not found, or if '`-B`' was not specified, the driver tries two standard prefixes, which are '`/usr/lib/gcc/`' and '`/usr/local/lib/gcc-lib/`'. If neither of those results in a file name that is found, the unmodified program name is searched for using the directories specified in your '`PATH`' environment variable.

'`-B`' prefixes that effectively specify directory names also apply to libraries in the linker, because the compiler translates these options into '`-L`' options for the linker. They also apply to includes files in the preprocessor, because the compiler translates these options into '`-isystem`' options for the preprocessor. In this case, the compiler appends '`include`' to the prefix.

The run-time support file '`libgcc.a`' can also be searched for using the '`-B`' prefix, if needed. If it is not found there, the two standard prefixes above are tried, and that is all. The file is left out of the link if it is not found by those means.

Another way to specify a prefix much like the '`-B`' prefix is to use the environment variable `GCC_EXEC_PREFIX`. See Section 2.17 [Environment Variables], page 67.

`-specs=`*file*

Process *file* after the compiler reads in the standard '`specs`' file, in order to override the defaults that the '`gcc`' driver program uses when determining what switches to pass to '`cc1`', '`cc1plus`', '`as`', '`ld`', etc. More than one '`-specs=`'*file* can be specified on the command line, and they are processed in order, from left to right.

2.13 Specifying subprocesses and the switches to pass to them

`GCC` is a driver program. It performs its job by invoking a sequence of other programs to do the work of compiling, assembling and linking. GCC interprets its command-line parameters and uses these to deduce which programs it should invoke, and which command-line options it ought to place on their command lines. This behaviour is controlled by *spec strings*. In most cases there is one spec string for each program that GCC can invoke, but a few programs have multiple spec strings to control their behaviour. The spec strings built into GCC can be overridden by using the '`-specs=`' command-line switch to specify a spec file.

Spec files are plaintext files that are used to construct spec strings. They consist of a sequence of directives separated by blank lines. The type of directive is determined by the first non-whitespace character on the line and it can be one of the following:

`%`*command* Issues a *command* to the spec file processor. The commands that can appear here are:

`%include <`*file*`>`

Search for *file* and insert its text at the current point in the specs file.

`%include_noerr <`*file*`>`

Just like '`%include`', but do not generate an error message if the include file cannot be found.

`%rename` *old_name new_name*

Rename the spec string *old_name* to *new_name*.

`*[`*spec_name*`]:`

This tells the compiler to create, override or delete the named spec string. All lines after this directive up to the next directive or blank line are considered to be the text for the spec string. If this results in an empty string then the spec will be deleted. (Or, if the spec did not exist, then nothing will happened.) Otherwise, if the spec does not currently exist a new spec will be created. If the spec does exist then its contents will be overridden by the text of this directive, unless the first character of that text is the '`+`' character, in which case the text will be appended to the spec.

`[`*suffix*`]:` Creates a new '`[`*suffix*`] spec`' pair. All lines after this directive and up to the next directive or blank line are considered to make up the spec string for the indicated suffix. When the compiler encounters an input file with the named suffix, it will processes the spec string in order to work out how to compile that file. For example:

```
.zz:
z-compile -input %i
```

This says that any input file whose name ends in '`.zz`' should be passed to the program '`z-compile`', which should be invoked with the command-line switch '`-input`' and with the result of performing the '`%i`' substitution. (See below.)

As an alternative to providing a spec string, the text that follows a suffix directive can be one of the following:

`@`*language* This says that the suffix is an alias for a known *language*. This is similar to using the `-x` command-line switch to GCC to specify a language explicitly. For example:

```
.ZZ:
@c++
```

Says that .ZZ files are, in fact, C++ source files.

`#`*name* This causes an error messages saying:

name `compiler not installed on this system.`

GCC already has an extensive list of suffixes built into it. This directive will add an entry to the end of the list of suffixes, but since the list is searched from the end backwards, it is effectively possible to override earlier entries using this technique.

GCC has the following spec strings built into it. Spec files can override these strings or create their own. Note that individual targets can also add their own spec strings to this list.

```
asm          Options to pass to the assembler
asm_final    Options to pass to the assembler post-processor
cpp          Options to pass to the C preprocessor
cc1          Options to pass to the C compiler
cc1plus      Options to pass to the C++ compiler
endfile      Object files to include at the end of the link
link         Options to pass to the linker
lib          Libraries to include on the command line to the linker
libgcc       Decides which GCC support library to pass to the linker
linker       Sets the name of the linker
predefines   Defines to be passed to the C preprocessor
signed_char  Defines to pass to CPP to say whether char is signed by default
startfile    Object files to include at the start of the link
```

Here is a small example of a spec file:

```
%rename lib                 old_lib

*lib:
--start-group -lgcc -lc -leval1 --end-group %(old_lib)
```

This example renames the spec called '`lib`' to '`old_lib`' and then overrides the previous definition of '`lib`' with a new one. The new definition adds in some extra command-line options before including the text of the old definition.

Spec strings are a list of command-line options to be passed to their corresponding program. In addition, the spec strings can contain '`%`'-prefixed sequences to substitute variable text or to conditionally insert text into the command line. Using these constructs it is possible to generate quite complex command lines.

Here is a table of all defined '`%`'-sequences for spec strings. Note that spaces are not generated automatically around the results of expanding these sequences. Therefore you can concatenate them together or combine them with constant text in a single argument.

`%%` Substitute one '`%`' into the program name or argument.

`%i` Substitute the name of the input file being processed.

`%b` Substitute the basename of the input file being processed. This is the substring up to (and not including) the last period and not including the directory.

`%d` Marks the argument containing or following the '`%d`' as a temporary file name, so that that file will be deleted if GCC exits successfully. Unlike '`%g`', this contributes no text to the argument.

`%g`*suffix* Substitute a file name that has suffix *suffix* and is chosen once per compilation, and mark the argument in the same way as '`%d`'. To reduce exposure to denial-of-service attacks, the file name is now chosen in a way that is hard to predict even when previously chosen file names are known. For example, '`%g.s ... %g.o ... %g.s`' might turn into '`ccUVUUAU.s ccXYAXZ12.o ccUVUUAU.s`'. *suffix* matches the regexp '`[.A-Za-z]*`' or the special string '`%O`', which is treated exactly as if '`%O`' had been preprocessed. Previously, '`%g`' was simply substituted with a file name chosen once per compilation, without regard to any appended suffix (which was therefore treated just like ordinary text), making such attacks more likely to succeed.

`%u`*suffix* Like '`%g`', but generates a new temporary file name even if '`%u`*suffix*' was already seen.

`%U`*suffix* Substitutes the last file name generated with '`%u`*suffix*', generating a new one if there is no such last file name. In the absence of any '`%u`*suffix*', this is just like '`%g`*suffix*', except they don't share the same suffix *space*, so '`%g.s ... %U.s ... %g.s ... %U.s`' would involve the generation of two distinct file names, one for each '`%g.s`' and another for each '`%U.s`'. Previously, '`%U`' was simply substituted with a file name chosen for the previous '`%u`', without regard to any appended suffix.

`%w` Marks the argument containing or following the '`%w`' as the designated output file of this compilation. This puts the argument into the sequence of arguments that '`%o`' will substitute later.

`%o` Substitutes the names of all the output files, with spaces automatically placed around them. You should write spaces around the '`%o`' as well or the results are undefined. '`%o`' is for use in the specs for running the linker. Input files whose names have no recognized suffix are not compiled at all, but they are included among the output files, so they will be linked.

`%O` Substitutes the suffix for object files. Note that this is handled specially when it immediately follows '`%g, %u, or %U`', because of the need for those to form complete file names. The handling is such that '`%O`' is treated exactly as if it had already been substituted, except that '`%g, %u, and %U`' do not currently support additional *suffix* characters following '`%O`' as they would following, for example, '`.o`'.

`%p` Substitutes the standard macro predefinitions for the current target machine. Use this when running `cpp`.

`%P` Like '`%p`', but puts '`__`' before and after the name of each predefined macro, except for macros that start with '`__`' or with '*_L*', where *L* is an uppercase letter. This is for ANSI C.

`%I` Substitute a '`-iprefix`' option made from GCC_EXEC_PREFIX.

`%s` Current argument is the name of a library or startup file of some sort. Search for that file in a standard list of directories and substitute the full name found.

`%e`*str* Print *str* as an error message. *str* is terminated by a newline. Use this when inconsistent options are detected.

`%|` Output '`-`' if the input for the current command is coming from a pipe.

`%(`*name*`)` Substitute the contents of spec string *name* at this point.

`%[`*name*`]` Like '`%(...)`' but put '`__`' around '`-D`' arguments.

`%x{`*option*`}` Accumulate an option for '`%X`'.

`%X` Output the accumulated linker options specified by '`-Wl`' or a '`%x`' spec string.

`%Y` Output the accumulated assembler options specified by '`-Wa`'.

`%Z` Output the accumulated preprocessor options specified by '`-Wp`'.

`%v1` Substitute the major version number of GCC. (For version 2.9.5, this is 2.)

`%v2` Substitute the minor version number of GCC. (For version 2.9.5, this is 9.)

`%a` Process the `asm` spec. This is used to compute the switches to be passed to the assembler.

`%A` Process the `asm_final` spec. This is a spec string for passing switches to an assembler post-processor, if such a program is needed.

`%l` Process the `link` spec. This is the spec for computing the command line passed to the linker. Typically it will make use of the '`%L %G %S %D and %E`' sequences.

`%D` Dump out a '`-L`' option for each directory that GCC believes might contain startup files. If the target supports multilibs then the current multilib directory will be prepended to each of these paths.

`%L` Process the `lib` spec. This is a spec string for deciding which libraries should be included on the command line to the linker.

`%G` Process the `libgcc` spec. This is a spec string for deciding which GCC support library should be included on the command line to the linker.

`%S` Process the `startfile` spec. This is a spec for deciding which object files should be the first ones passed to the linker. Typically this might be a file named '`crt0.o`'.

`%E` Process the `endfile` spec. This is a spec string that specifies the last object files that will be passed to the linker.

`%C` Process the `cpp` spec. This is used to construct the arguments to be passed to the C preprocessor.

`%c` Process the `signed_char` spec. This is intended to be used to tell cpp whether a char is signed. It typically has the definition:

```
%{funsigned-char:-D__CHAR_UNSIGNED__}
```

`%1` Process the `cc1` spec. This is used to construct the options to be passed to the actual C compiler ('`cc1`').

`%2` Process the `cc1plus` spec. This is used to construct the options to be passed to the actual C++ compiler (‘`cc1plus`’).

`%*` Substitute the variable part of a matched option. See below. Note that each comma in the substituted string is replaced by a single space.

`%{S}` Substitutes the `-S` switch, if that switch was given to GCC. If that switch was not specified, this substitutes nothing. Note that the leading dash is omitted when specifying this option, and it is automatically inserted if the substitution is performed. Thus the spec string ‘`%{foo}`’ would match the command-line option ‘`-foo`’ and would output the command line option ‘`-foo`’.

`%W{S}` Like %`{S}` but mark last argument supplied within as a file to be deleted on failure.

`%{S*}` Substitutes all the switches specified to GCC whose names start with `-S`, but which also take an argument. This is used for switches like ‘`-o, -D, -I`’, etc. GCC considers ‘`-o foo`’ as being one switch whose names starts with ‘`o`’. %`{o*}` would substitute this text, including the space. Thus two arguments would be generated.

`%{^S*}` Like %`{S*}`, but don’t put a blank between a switch and its argument. Thus %`{^o*}` would only generate one argument, not two.

`%{<S}` Remove all occurences of `S` from the command line. Note - this command is position dependent. ‘`%`’ commands in the spec string before this option will see `S`, ‘`%`’ commands in the spec string after this option will not.

`%{S*:X}` Substitutes `X` if one or more switches whose names start with `-S` are specified to GCC. Note that the tail part of the `-S` option (i.e. the part matched by the ‘`*`’) will be substituted for each occurrence of ‘`%*`’ within `X`.

`%{S:X}` Substitutes `X`, but only if the ‘`-S`’ switch was given to GCC.

`%{!S:X}` Substitutes `X`, but only if the ‘`-S`’ switch was *not* given to GCC.

`%{|S:X}` Like %`{S:X}`, but if no `S` switch, substitute ‘`-`’.

`%{|!S:X}` Like %`{!S:X}`, but if there is an `S` switch, substitute ‘`-`’.

`%{.S:X}` Substitutes `X`, but only if processing a file with suffix `S`.

`%{!.S:X}` Substitutes `X`, but only if *not* processing a file with suffix `S`.

`%{S|P:X}` Substitutes `X` if either `-S` or `-P` was given to GCC. This may be combined with ‘`!`’ and ‘`.`’ sequences as well, although they have a stronger binding than the ‘`|`’. For example a spec string like this:

```
%{.c:-foo} %{!.c:-bar} %{.c|d:-baz} %{!.c|d:-boggle}
```

will output the following command-line options from the following input command-line options:

```
fred.c        -foo -baz
jim.d         -bar -boggle
-d fred.c     -foo -baz -boggle
-d jim.d      -bar -baz -boggle
```

The conditional text `X` in a %`{S:X}` or %`{!S:X}` construct may contain other nested ‘`%`’ constructs or spaces, or even newlines. They are processed as usual, as described above.

The ‘-O, -f, -m, and -W’ switches are handled specifically in these constructs. If another value of ‘-O’ or the negated form of a ‘-f, -m, or -W’ switch is found later in the command line, the earlier switch value is ignored, except with {S*} where S is just one letter, which passes all matching options.

The character ‘|’ at the beginning of the predicate text is used to indicate that a command should be piped to the following command, but only if ‘-pipe’ is specified.

It is built into GCC which switches take arguments and which do not. (You might think it would be useful to generalize this to allow each compiler’s spec to say which switches take arguments. But this cannot be done in a consistent fashion. GCC cannot even decide which input files have been specified without knowing which switches take arguments, and it must know which input files to compile in order to tell which compilers to run).

GCC also knows implicitly that arguments starting in ‘-l’ are to be treated as compiler output files, and passed to the linker in their proper position among the other output files.

2.14 Specifying Target Machine and Compiler Version

Some GNU C installations allow you to use the same command name to compile for many different targets – one of them perhaps native applications on your workstation. To manage support, MIPS SDE does not provide this level of option; if you have a different GNU compiler, it will be called ‘gcc’ not ‘sde-gcc’. MIPS SDE does allow you to select from a large choice of MIPS cross-compilation targets; but that’s described in the next section, Section 2.15 [Different CPUs and Configurations], page 52.

2.15 Different CPUs and Configurations

The compiler provided with MIPS SDE can generate code for a wide range of MIPS CPUs. Many of these CPU types have a special option, starting with ‘-m’, which configures the compilation process towards that CPU’s particular features–for example, R4640 vs R3000, floating-point coprocessor or none. A single installed version of the compiler can compile for any model or configuration, according to the options specified.

These ‘-m’ options are defined for the MIPS family of computers:

`-mcpu=`*cpu_type*

Set up the compiler for a particular type or family of CPU. This affects instruction scheduling where the compiler knows about the CPU’s timing habits, and produces usually-correct defaults for some of the other options defined below. What it doesn’t do is to pick the major instruction set variants, which are set explicitly with options like ‘-mips3’. Unless you specify otherwise, the compiler will only use the “MIPS I” instruction set of the earliest MIPS CPUs.

Legitimate values for *cpu_type* are:

`r1900, pr1900`
: Toshiba TX19 CPU core, like ‘r3900’ but with MIPS16.

`r2k, r2000`

`r3k, r3000`
: Standard MIPS I CPUs.

`r3900, pr3900, tx3900`
: Philips/Toshiba TX39 MIPS I CPU core, with fast multiply-accumulate and branch-likely.

`rc32364, rc323xx`
: IDT 32-bit CPU core – like ‘`r3000`’, but with MIPS32-style extensions.

`r4k, r4000, r4400`
: Standard MIPS III CPUs (long pipeline).

`4kc, 4km, 4kec, 4kem`
: MIPS Technologies 4K range of synthesisable MIPS32 and MIPS32 release 2 core CPUs.

`4kp, 4kep` MIPS Technologies 4K synthesisable core CPUs with slow multiplier.

`4ksc, 4ksd` MIPS Technologies 4KSc and 4KSd synthesisable MIPS32 core CPUs for smartcards, with SmartMIPS extensions.

`m4k` MIPS Technologies M4K compact, synthesisable MIPS32 release 2 core CPU.

`5kc, 5kf` MIPS Technologies 5K range of synthesisable MIPS64 core CPUs.

`24kc, 24kf` MIPS Technologies 24K range of high-performance, synthesisable MIPS32 release 2 core CPUs.

`20kc` MIPS Technologies 20Kc dual-issue MIPS64 hard core CPU.

`25kf` MIPS Technologies 25Kf dual-issue MIPS64 hard core CPU.

`cw400x, cw4001, cw4002, cw4003`
: LSI miniRISC CPU cores from the CW400x/MR400x families; simple pipeline, LSI-style multiply-accumulate.

`cw401x, cw4010, cw4011`
: LSI CW401x/MR401x family, superscalar miniRISC CPU cores with dual issue and some additional instructions.

`atm2, atmizer2, apu`
: LSI ATMizer-2, a specialised Cw401x CPU for ATM.

`tr4101, ev4101`
: LSI TinyRisc CPU core – like ‘`cw400x`’, but with MIPS16.

`r4200, vr4200`
: Standard MIPS III CPUs, but with short pipeline and shared integer/floating-point ALU.

`r4100, vr4100`
: NEC Vr410x MIPS III core – like ‘`vr4200`’ but no floating-point.

`r4111, vr4111`
: NEC Vr4111 core – like ‘`vr4100`’ but with MIPS16.

`r4300, vr4300`
: NEC Vr4300 MIPS III CPU – like ‘`vr4200`’, but faster.

`r4320, vr4320`
: NEC Vr4300 variant, with integer multiply-accumulate.

`r4600, r4700`
: IDT MIPS III CPUs, with short pipeline.

`r4640, r4650`
: IDT R4600 variants, with integer multiply-accumulate, but only single-precision floating-point.

`r5k, r5000`
: R5000 and family, standard MIPS IV CPUs, limited dual-issue of integer and floating-point.

`rm52xx, rm5230, rm5260, rm5270`
: PMC-Sierra RM52xx CPU – like '`r5000`' but with integer multiply-accumulate instructions.

`r54xx, r5400, r5432, r5464`
`vr54xx, vr5400, vr5432, vr5464`
: NEC Vr54xx CPU – like '`r5000`' but with integer multiply-accumulate, rotate, multimedia extensions, dual-issue and non-blocking loads.

`rc64574, rc64575, rc6457x`
: IDT 64-bit CPUs – like '`r5000`', but with MIPS32-style extensions.

`rm7k, rm7000`
: PMC-Sierra RM70xx CPU – like '`rm52xx`' but with dual-issue and non-blocking loads.

`r6k, r6000`
: Ancient MIPS II CPU.

`r10k, r10000`
: Top of the range MIPS IV CPU, with out-of-order execution.

`-mips1, -mips2, -mips3, -mips4`
: The MIPS instruction set has been extended several times, and each major extension is a superset of the one before. Each option tells the compiler to make use of instructions from the appropriate instruction set subset; the default is '`-mips1`' which uses only the MIPS I instruction set compatible with every MIPS CPU there ever was.

  Each instruction set variant has a different default *cpu_type*; they are '`r3000`', '`r6000`', '`r4000`' and '`r5000`' respectively.

  Although the '`-mips2`' default CPU is long obsolete, it's still a useful option to build code for a MIPS III CPU when you don't need any 64-bit specific instructions – roughly equivalent to '`-mips3 -mgp32 -mfp32`', but your exception handlers don't have to save and restore the full 64-bit register set.

`-mips5`
: Treated identically to '`-mips4`' by the compiler, which currently has no support for the paired-single vector floating-point format which this instruction set introduced.

`-mips32`
: MIPS Technologies' rationalisation of the 32-bit MIPS instruction set, which at the application level looks a lot like MIPS II with the MIPS IV integer extensions (conditional move, etc), plus multiply-add, multiply-subtract, three-operand multiply, count leading zeroes and ones. This instruction set is likely to be found on an increasing range of 32-bit MIPS-based CPUs and cores, from a number of vendors.

`-mips64` MIPS Technologies' rationalisation of the 64-bit MIPS instruction set, which at the application level looks a lot like MIPS IV, plus multiply-add, multiply-subtract, three-operand multiply, count leading zeroes and ones. This instruction set is likely to be found on an increasing range of 64-bit MIPS-based CPUs and cores, from a number of vendors.

`-mips32r2, -mips64r2`
The release 2 updates to the MIPS32 and MIPS64 architecture add some useful new instructions, including: bit-rotate, bit-field insert/extract, byte swapping, and register sign-extend.

`-mips16` MIPS16 is an extension to the base instruction set which provides a subset of true 32-bit MIPS instructions, with a restricted set of registers, and coded as 16-bit instructions. It can make for much smaller program binaries. CPUs supporting MIPS16 switch from interpreting conventional 32-bit MIPS instructions to MIPS16 instructions when they are asked to fetch an instruction from an odd address.

All functions in a C or C++ module built with '`-mips16`' are compiled using the narrower instructions, unless modified by the '`nomips16`' function attribute (see Section 3.23 [Function Attributes], page 86). MIPS16 functions may be freely interlinked with true MIPS functions. It is difficult and usually pointless to try to write assembler code for MIPS16; but inside an assembler module you can mix MIPS16 and conventional code using '`.set mips16`' and '`.set nomips16`' directives.

MIPS16 CPUs can be either 32-bit or 64-bit implementations. To generate 64-bit MIPS16 code, you should specify a 64-bit base ISA, e.g '`-mips3 -mips16`'.

`-mips16e` MIPS16e is an enhancement of the MIPS16 instruction set, which provides even greater code size reductions. It is only ever available with a MIPS32 or MIPS64 compliant CPU.

Strictly the '`-mips16e`' option is not required—using the '`-mips16`' option will generate MIPS16e code if applied as a modifier to '`-mips32`' or '`-mips64`'. The only difference is that a '`-mips16e`' option used on its own will generate MIPS32 code in functions marked with the '`nomips16`' attribute, whereas '`-mips16`' would generate MIPS I code.

`-msmartmips`
SmartMIPS is an extension to the MIPS32 instruction set which provides a number of new instructions which target smartcard and cryptographic applications. The compiler will make use of its rotate and indexed load-word instructions, and the other SmartMIPS instructions can be used in assembler code or C `asm` expressions.

`-mips3D` MIPS-3D is an extension to the MIPS64 instruction set, and is treated identically to '`-mips64`' by the compiler. The new floating-point instructions can be used in assembler code or C `asm` expressions.

`-mcode-xonly`
For MIPS16 code generation only, this flag forces all string constants and jump tables to be placed in the read-only data section, rather than the text/code section. This option is required if you are compiling code for a CPU and virtual memory operating system which mark code pages as "execute only", or a CPU which has separate instruction and data memories. It may result in somewhat larger and slower code. Note that

this option does not disable the MIPS16 `lwpc` instruction, which is still used to load 32-bit immediates from the text section, following the current function—it is handled specially by the CPU, which treats the pc-relative data load as an instruction fetch.

`-mno-data-in-code`
: In MIPS16 code the compiler normally places implicit constants and strings (but not C variables declared with the `const` attribute) in the '`text`' section, immediately following the referencing function, where they can be accessed using short PC-relative addresses.

  Some MIPS Technologies cores support independent instruction and data memories (SPRAM), which can't read data from the I-memory without special hardware support. Use the '`-mno-data-in-code`' option when compiling MIPS16 code for a CPU like that. It will produce *significantly* larger code; so don't use it unless you really need it. When this option is used, it switches off the '`-G 0`' override which is normally used for MIPS16 code, so that it can place the implicit constants in the small data section and access them via the *$gp* register. If you specify the '`-G 0`' option explicitly, then the generated code will get even larger.

`-muse-all-regs`
: For MIPS16 code generation only, this option instructs the compiler to make use of all the general purpose registers, and not just the eight which are directly accessible by MIPS16 code. This may generate better code in some cases, but should be treated as experimental. You probably only want to use this with MIPS16e code, and not the original MIPS16 instruction set, since storing and reloading the additional callee-saved registers in the function prologue/epilogue would be very expensive for vanilla MIPS16.

`-mbranch-likely`
: Allows the use of branch-likely instructions when generating MIPS32 or MIPS64 code. These instructions are officially deprecated and are not used by default when '`-mips32`' or '`-mips64`' is selected.

`-mfp32`
: Assume that 32 32-bit floating point registers are available, but only the even-numbered 16 are used for arithmetic (the odd-numbered registers are used quietly by the assembler for loading/storing the high-order bits of double-precision values). This is the default when a 32-bit ISA is specified or implied.

`-mfp64`
: Assume that 32 64-bit floating point registers are available. This is the default when a 64-bit ISA is specified or implied, and is usually illegal with 32-bit ISAs. The exception is that it can be used with '`-mips32r2`', since MIPS32 release 2 adds new instructions which allow a 32-bit CPU to be combined with a 64-bit FPU.

`-mgp32`
: Assume that the 32 general purpose registers are 32 bits wide. This is the default when a 32-bit ISA is specified or implied. Such code will run correctly on 64-bit MIPS CPUs so long as every function which might call your code is built the same way.

`-mgp64`
: Assume that the general purpose registers are 64 bits wide. This is the default when a 64-bit ISA is specified or implied, and illegal unless your CPU implements a 64-bit instruction set; so it's hard to see when you'd use this option explicitly.

`-mcheck-zero-division`
`-mno-check-zero-division`
`-mdiv-checks`
`-mno-div-checks`

Do, or don't, generate code to guarantee that integer division by zero will be detected. By default, detection is disabled for MIPS16, but enabled for the 32-bit instruction sets.

For MIPS CPUs the compiler generates code that explicitly checks for zero-valued divisors and traps when one is detected. Use of '`-mno-check-zero-division`' suppresses such checking.

The '`-mdiv-checks`' and '`-mno-div-checks`' flags are simply aliases for (respectively) the '`-mcheck-zero-division`' and '`-mno-check-zero-division`' flags, for compatibility with previous versions of MIPS SDE.

`-msplit-addresses`
`-mno-split-addresses`

Do, or don't, generate two separate assembler instructions to load the high and low parts of address constants. Defaults to on, which allows the compiler to schedule the two instructions separately and generate more optimal code.

`-mrnames`
`-mno-rnames`

The '`-mrnames`' switch says to output code using the MIPS software names for the registers, instead of the hardware names (ie, *$a0* instead of *$4*). On by default in MIPS SDE, because our version of the GNU assembler is happy with it, and it makes the compiler's assembler output easier to read.

`-mgpopt`
`-mno-gpopt`

The '`-mgpopt`' switch says to write all of the data declarations before the instructions in the text section, this allows the MIPS assembler to generate one word memory references instead of using two words for short global or static data items. This is on by default if optimization is selected.

`-mstats`
`-mno-stats`

For each non-inline function processed, the '`-mstats`' switch causes the compiler to emit one line to the standard error file to print statistics about the program (number of registers saved, stack size, etc.).

`-mmemcpy`
`-mno-memcpy`

The '`-mmemcpy`' switch makes all block moves call the appropriate string function ('`memcpy`' or '`bcopy`') instead of possibly generating inline code. It can reduce the size of your program.

`-msoft-float`
`-mhard-float`
`-mno-float`

Generate output using implicit library calls for floating point, so you can build for machines with no floating point hardware. The floating point emulation library is included in MIPS SDE. '`-mhard-float`' is the opposite, and it's the default, so rarely specified explicitly.

The '`-mno-float`' switch implies '`-msoft-float`' but goes further, by picking optional library variants at link time which don't provide any support for floating point data, and are therefore significantly smaller.

`-mslow-mul`

Fine-tune the compiler to particular LSI core CPUs (based on '`cw401x`') which can optionally use a smaller, slower integer multiplier.

`-mno-mul` Don't generate integer multiply/divide instructions at all—the compiler will instead insert calls to multiply/divide subroutines. You'll also need to build special libraries compiled with this option (they are not provided with MIPS SDE), or be prepared to catch the "undefined instruction" trap and emulate any multiply/divide instructions which get included from the libraries. The MIPS SDE run-time kit includes a sample instruction emulator.

`-mmad` Emit r4650-style `mad` and `mul` instructions. Normally the compiler decides whether it can use multiply-add and other extended multiplier features automatically, using the selected CPU type and ISA.

`-mconst-mult`

Encourages the compiler to use the hardware multiply instruction for all constant multiplications, if the multiply can be accomplished in fewer instructions than the equivalent operation performed inline with shifts and adds. This is the default for MIPS16 code, or if the '`-Os`' optimization level is selected. In all other cases the compiler will default to preferring a shift/add sequence if that would be faster (rather than smaller) than using the hardware multiplier.

`-mno-const-mult`

Forces the compiler to change its default for MIPS16 code, or when '`-Os`' is specified, and use speed rather than space as the constraint on whether to use the hardware multiply instructions.

`-membedded-data`
`-mno-embedded-data`
`-mgpconst`
`-mno-gpconst`

Allocate variables to the read-only data section first if possible, then next in the small data section if possible, otherwise in data. This gives slightly slower code than the default, but reduces the amount of RAM required when executing, and thus may be preferred for some embedded systems.

Note that '`-membedded-data`' is the default setting for MIPS SDE. Alternatively select '`-mno-embedded-data`' if you want small constants to be placed in the small data section, instead of the read-only data section, for increased speed.

The ‘`-mno-gpconst`’ and ‘`-mgpconst`’ flags are just aliases to the flags ‘`-membedded-data`’ and ‘`-mno-embedded-data`’ respectively, for compatibility with older versions of MIPS SDE.

`-muninit-const-in-rodata`
`-mno-uninit-const-in-rodata`
When used together with ‘`-membedded-data`’, this flag will cause uninitialized const variables to be placed in the read-only data section. The default action is to make uninitialised const variables “common”, which places them in the read-write data section.

`-mlong-calls`
`-mno-long-calls`
Do all calls with the `jalr` instruction, which requires loading a function’s address into a register before the call. You may need to use this switch if you link a program linked so that a calling function and its matching callee are in separate 256Mbyte segments of memory. But this is would be a bit drastic because it makes all calls much longer and slower. You could use the ‘`longcall`’ attribute (see Section 3.23 [Function Attributes], page 86) to mark particular functions known to require a long call.

`-msingle-float`
`-mdouble-float`
The ‘`-msingle-float`’ switch tells gcc to assume that the floating point coprocessor only supports single precision operations, as on IDT’s R4640 and R4650 CPUs. The ‘`-mdouble-float`’ switch permits gcc to use double precision operations, and this is the default, even if you specified ‘`-mcpu=r4650`’.

`-mcommon-prolog`
Use built-in subroutines to save/restore registers at function entry and exit, where this would make the code smaller. Can save a significant amount of space, at the cost of a small performance hit if you have lots of little functions. This doesn’t work with MIPS16, for which you must use the ‘`-mentry`’ switch below.

`-mentry` Used only with ‘`-mips16`’, this compresses function entry and exit code by encoding common register save/restore actions into fictional entry/exit instructions. The non-existent instruction codes produce an “unimplemented instruction” trap, and a system using this relies on a trap handler to implement these instructions. The MIPS SDE run-time kit includes a sample entry/exit trap handler.

`-EL` Compile code for the processor in little endian mode. The requisite libraries are assumed to exist.

`-EB` Compile code for the processor in big endian mode. The requisite libraries are assumed to exist.

`-G` *num* Put global and static items less than or equal to *num* bytes into special “small data” or “small bss” sections instead of the normal data or bss section. In co-operation with a run-time system which maintains the global pointer register (*gp* or *$28*) to point to the accumulated small data items, this allows the assembler to emit one word instruction for variable load/stores.

By default, *num* is 8, but if your memory map or run-time system will not permit this technique then use ‘`-G 0`’ to disable the optimisation. You can also increase the value

to place more variables into the "fast" sections, but take care not to overflow the 64KB maximum size of the small data region.

The '`-G` *num*' switch is also passed to the assembler and linker. All modules must be compiled with the same '`-G` *num*' value.

Because GNU C is a mighty undertaking there are options supported by the compiler, but not supported by the rest of the MIPS SDE toolkit. For completeness they are listed here:

`-mabi=32`
`-mabi=o64`
`-mabi=n32`
`-mabi=64`
`-mabi=eabi`
`-mabi=meabi`

Generate code for the indicated ABI. This option should be considered experimental in MIPS SDE. The default is '`-mabi=32`' in all cases, even when a 64-bit ISA has been selected, and all of the supplied MIPS SDE libraries have been built with the 32-bit ABI.

Note that '`-mabi=o64`' is a non-standard, undocumented ABI used by some other GNU toolchains for 64-bit code—essentially it's what you get if you stretch '`-mabi=32`' so that all argument registers and stack parameters are 64-bits wide. We only provide '`o64`' mode for compatibility with customers with code based that assume that calling convention, but it won't work with the MIPS SDE libraries, and is deprecated.

`-mint64` Force long, int and pointer types to be 64 bits wide. This is effective only if a 64-bit ISA is also specified. It is incompatible with the supplied MIPS SDE header files and libraries.

`-mlong64` Force long and pointer types to be 64 bits wide, but the int type remains 32 bits wide. This is effective only if a 64-bit ISA is also specified. It is incompatible with the supplied MIPS SDE header files and libraries.

`-mlong32` Force long, int, and pointer types to be 32 bits wide.

If none of '`-mlong32`', '`-mlong64`', or '`-mint64`' are set, then the size of ints, longs, and pointers depends on the ABI and ISA chosen. For '`-mabi=32`', and '`-mabi=n32`', ints and longs are 32 bits wide. For '`-mabi=64`', ints are 32 bits, and longs are 64 bits wide. For '`-mabi=eabi`' and '`-mabi=meabi`' ints are always 32 bits, but the size of longs tracks the register size of the selected ISA. The width of pointers is always the same as the width of longs.

`-mgas`
`-mmips-as` Generate code for the GNU assembler (the default), or for the old MIPS Computers proprietary assembler. Selecting the old MIPS Computers assembler is unlikely to work well with the MIPS SDE toolchain, which only includes the GNU assembler.

`-mhalf-pic`
`-mno-half-pic`

Put pointers to extern references into the data section and load them up, rather than putting the references in the text section. Added to the compiler only for compiling OSF/1 applications. This is unlikely to work correctly with the MIPS SDE toolchain.

`-membedded-pic`
`-mno-embedded-pic`
: Generate PIC code suitable for some embedded systems. All calls are made using PC relative address, and all data is addressed using the *gp* register. No more than 65536 bytes of global data may be used. This requires GNU as and GNU ld which do most of the work. This currently only works on targets which use ECOFF; it does not work with ELF. MIPS SDE uses ELF, so this feature is not supported.

`-mabicalls`
`-mno-abicalls`
: Emit (or do not emit) the pseudo operations '`.abicalls`', '`.cpload`', and '`.cprestore`' that some Unix systems (such as IRIX, Linux and LynxOS) use for position independent code. In MIPS SDE this option defaults to disabled, but can be enabled if you really know what you are doing.

`-mmips-tfile`
`-mno-mips-tfile`
: These options relate to the old MIPS ECOFF object code format, and are not supported by MIPS SDE.

2.16 Options for Code Generation Conventions

These machine-independent options control the interface conventions used in code generation.

Most of them have both positive and negative forms; the negative form of '`-ffoo`' would be '`-fno-foo`'. In the table below, only one of the forms is listed—the one which is not the default. You can figure out the other form by either removing '`no-`' or adding it.

This level of control is not really compatible with use of a bundled system like MIPS SDE. Some of the options documented here will break compatibility with the MIPS SDE libraries or other parts of the toolchain.

`-fexceptions`
: Enable exception handling. Generates extra code needed to propagate exceptions. For some targets, this implies GNU CC will generate frame unwind information for all functions, which can produce significant data size overhead, although it does not affect execution. If you do not specify this option, GNU CC will enable it by default for languages like C++ which normally require exception handling, and disable it for languages like C that do not normally require it. However, you may need to enable this option when compiling C code that needs to interoperate properly with exception handlers written in C++. You may also wish to disable this option if you are compiling older C++ programs that don't use exception handling.

`-funwind-tables`
: Similar to `-fexceptions`, except that it will just generate any needed static data, but will not affect the generated code in any other way. You will normally not enable this option; instead, a language processor that needs this handling would enable it on your behalf.

`-fpcc-struct-return`
: Return "short" `struct` and `union` values in memory like longer ones, rather than in registers. This convention is less efficient, but it has the advantage of allowing intercallability between GCC-compiled files and files compiled with other compilers.

  The precise convention for returning structures in memory depends on the target configuration macros.

  Short structures and unions are those whose size and alignment match that of some integer type.

`-freg-struct-return`
: Use the convention that `struct` and `union` values are returned in registers when possible. This is more efficient for small structures than '`-fpcc-struct-return`'.

  If you specify neither '`-fpcc-struct-return`' nor its contrary '`-freg-struct-return`', GCC defaults to whichever convention is standard for the target. If there is no standard convention, GCC defaults to '`-fpcc-struct-return`', except on targets where GCC is the principal compiler. In those cases, we can choose the standard, and we chose the more efficient register return alternative.

`-fshort-enums`
: Allocate to an `enum` type only as many bytes as it needs for the declared range of possible values. Specifically, the `enum` type will be equivalent to the smallest integer type which has enough room.

`-fshort-double`
: Use the same size for `double` as for `float`.

`-fno-common`
: Allocate even uninitialized global variables in the data section of the object file, rather than generating them as common blocks. This has the effect that if the same variable is declared (without `extern`) in two different compilations, you will get an error when you link them.

  This setting is the default for MIPS16 code generation, since it allows more efficient access to locally defined variables. If that breaks your code then you can either use the "common" attribute on individual variables (see Section 3.29 [Variable Attributes], page 92), or use the positive '`-fcommon`' flag to override this globally.

`-fno-ident`
: Ignore the '`#ident`' directive.

`-fno-gnu-linker`
: Do not output global initializations (such as C++ constructors and destructors) in the form used by the GNU linker (on systems where the GNU linker is the standard method of handling them). Use this option when you want to use a non-GNU linker, which also requires using the `collect2` program to make sure the system linker includes constructors and destructors. (`collect2` is included in the GCC distribution.) For systems which *must* use `collect2`, the compiler driver `gcc` is configured to do this automatically.

`-finhibit-size-directive`
: Don't output a `.size` assembler directive, or anything else that would cause trouble if the function is split in the middle, and the two halves are placed at locations far apart

in memory. This option is used when compiling '`crtstuff.c`'; you should not need to use it for anything else.

`-fverbose-asm`
Put extra commentary information in the generated assembly code to make it more readable. This option is generally only of use to those who actually need to read the generated assembly code (perhaps while debugging the compiler itself).

'`-fno-verbose-asm`', the default, causes the extra information to be omitted and is useful when comparing two assembler files.

`-fvolatile`
Consider all memory references through pointers to be volatile.

`-fvolatile-global`
Consider all memory references to extern and global data items to be volatile. GCC does not consider static data items to be volatile because of this switch.

`-fvolatile-static`
Consider all memory references to static data to be volatile.

`-fpic` Generate position-independent code (PIC) suitable for use in a shared library, if supported for the target machine. Such code accesses all constant addresses through a global offset table (GOT). The dynamic loader resolves the GOT entries when the program starts (the dynamic loader is not part of GCC; it is part of the operating system). If the GOT size for the linked executable exceeds a machine-specific maximum size, you get an error message from the linker indicating that '`-fpic`' does not work; in that case, recompile with '`-fPIC`' instead. (These maximums are 16k on the m88k, 8k on the Sparc, and 32k on the m68k, RS/6000 and MIPS. The 386 has no such limit.)

Position-independent code requires special support, and therefore works only on certain machines and operating systems.

`-fPIC` If supported for the target machine, emit position-independent code, suitable for dynamic linking and avoiding any limit on the size of the global offset table. This option makes a difference on the m68k, m88k, and the Sparc.

Position-independent code requires special support, and therefore works only on certain machines and operating systems.

`-ffixed-`*reg*
Treat the register named *reg* as a fixed register; generated code should never refer to it (except perhaps as a stack pointer, frame pointer or in some other fixed role).

reg must be the name of a register. The register names accepted are machine-specific. For MIPS they are the register number, optionally preceded by a '$' (e.g. `$4`), or the software name of the register, again optionally preceded by a '$' (e.g. `$a0`).

This flag does not have a negative form, because it specifies a three-way choice.

`-fcall-used-`*reg*
Treat the register named *reg* as an allocable register that is clobbered by function calls. It may be allocated for temporaries or variables that do not live across a call. Functions compiled this way will not save and restore the register *reg*.

It is an error to used this flag with the frame pointer or stack pointer. Use of this flag for other registers that have fixed pervasive roles in the machine's execution model will produce disastrous results.

This flag does not have a negative form, because it specifies a three-way choice.

`-fcall-saved-`*reg*

Treat the register named *reg* as an allocable register saved by functions. It may be allocated even for temporaries or variables that live across a call. Functions compiled this way will save and restore the register *reg* if they use it.

It is an error to used this flag with the frame pointer or stack pointer. Use of this flag for other registers that have fixed pervasive roles in the machine's execution model will produce disastrous results.

A different sort of disaster will result from the use of this flag for a register in which function values may be returned.

This flag does not have a negative form, because it specifies a three-way choice.

`-fpack-struct`

Pack all structure members together without holes. Usually you would not want to use this option, since it makes the code suboptimal, and the offsets of structure members won't agree with system libraries.

`-fcheck-memory-usage`

Generate extra code to check each memory access. GCC will generate code that is suitable for a detector of bad memory accesses such as '`Checker`'.

Normally, you should compile all, or none, of your code with this option.

If you do mix code compiled with and without this option, you must ensure that all code that has side effects and that is called by code compiled with this option is, itself, compiled with this option. If you do not, you might get erroneous messages from the detector.

If you use functions from a library that have side-effects (such as `read`), you might not be able to recompile the library and specify this option. In that case, you can enable the '`-fprefix-function-name`' option, which requests GCC to encapsulate your code and make other functions look as if they were compiled with '`-fcheck-memory-usage`'. This is done by calling "stubs", which are provided by the detector. If you cannot find or build stubs for every function you call, you might have to specify '`-fcheck-memory-usage`' without '`-fprefix-function-name`'.

If you specify this option, you can not use the `asm` or `__asm__` keywords in functions with memory checking enabled. GNU CC cannot understand what the `asm` statement may do, and therefore cannot generate the appropriate code, so it will reject it. However, if you specify the function attribute `no_check_memory_usage` (see see Section 3.23 [Function Attributes], page 86, GNU CC will disable memory checking within a function; you may use `asm` statements inside such functions. You may have an inline expansion of a non-checked function within a checked function; in that case GNU CC will not generate checks for the inlined function's memory accesses.

If you move your `asm` statements to non-checked inline functions and they do access memory, you can add calls to the support code in your inline function, to indicate any

reads, writes, or copies being done. These calls would be similar to those done in the stubs described above.

`-fprefix-function-name`

Request GCC to add a prefix to the symbols generated for function names. GCC adds a prefix to the names of functions defined as well as functions called. Code compiled with this option and code compiled without the option can't be linked together, unless stubs are used.

If you compile the following code with '`-fprefix-function-name`'

```
extern void bar (int);
void
foo (int a)
{
  return bar (a + 5);
}
```

GCC will compile the code as if it was written:

```
extern void prefix_bar (int);
void
prefix_foo (int a)
{
  return prefix_bar (a + 5);
}
```

This option is designed to be used with '`-fcheck-memory-usage`'.

`-finstrument-functions`

Generate instrumentation calls for entry and exit to functions. Just after function entry and just before function exit, the following profiling functions will be called with the address of the current function and its call site. (On some platforms, `__builtin_return_address` does not work beyond the current function, so the call site information may not be available to the profiling functions otherwise.)

```
void __cyg_profile_func_enter (void *this_fn, void *call_site);
void __cyg_profile_func_exit  (void *this_fn, void *call_site);
```

The first argument is the address of the start of the current function, which may be looked up exactly in the symbol table.

This instrumentation is also done for functions expanded inline in other functions. The profiling calls will indicate where, conceptually, the inline function is entered and exited. This means that addressable versions of such functions must be available. If all your uses of a function are expanded inline, this may mean an additional expansion of code size. If you use '`extern inline`' in your C code, an addressable version of such functions must be provided. (This is normally the case anyways, but if you get lucky and the optimizer always expands the functions inline, you might have gotten away without providing static copies.)

A function may be given the attribute `no_instrument_function`, in which case this instrumentation will not be done. This can be used, for example, for the profiling functions listed above, high-priority interrupt routines, and any functions from which the profiling functions cannot safely be called (perhaps signal handlers, if the profiling routines generate output or allocate memory).

`-fstack-check`

Generate code to verify that you do not go beyond the boundary of the stack. You should specify this flag if you are running in an environment with multiple threads, but only rarely need to specify it in a single-threaded environment since stack overflow is automatically detected on nearly all systems if there is only one stack.

Note that this switch does not actually cause checking to be done; the operating system must do that. The switch causes generation of code to ensure that the operating system sees the stack being extended.

`-fstack-limit-register=`*reg*

`-fstack-limit-symbol=`*sym*

`-fno-stack-limit`

Generate code to ensure that the stack does not grow beyond a certain value, either the value of a register or the address of a symbol. If the stack would grow beyond the value, a signal is raised. For most targets, the signal is raised before the stack overruns the boundary, so it is possible to catch the signal without taking special precautions.

For instance, if the stack starts at address '`0x80000000`' and grows downwards you can use the flags '`-fstack-limit-symbol=__stack_limit`' and '`-Wl,--defsym,__stack_limit=0x7ffe0000`' which will enforce a stack limit of 128K.

`-funaligned-pointers`

Assume that all pointers contain unaligned addresses. On machines where unaligned memory accesses trap, this will result in much larger and slower code for all pointer dereferences, but the code will work even if addresses are unaligned.

`-funaligned-struct-hack`

Always access structure fields using loads and stores of the declared size. This option is useful for code that derefences pointers to unaligned structures, but only accesses fields that are themselves aligned. Without this option, gcc may try to use a memory access larger than the field. This might give an unaligned access fault on some hardware.

This option makes some invalid code work at the expense of disabling some optimizations. It is strongly recommended that this option not be used.

`-foptimize-comparisons`

Optimize multiple comparisons better within `&&` and `||` expressions. This is an experimental option. In some cases it can result in worse code. It depends on many factors. Now it is known only that the optimization works well for PPC740 and PPC750. This option switches on the following transformations:

```
(a != 0 || b != 0)  =>  ((a | b) != 0)
(a == 0 && b == 0)  =>  ((a | b) == 0)
(a != b || c != d)  =>  (((a ^ b) | (c ^ d)) != 0)
(a == b && c == d)  =>  (((a ^ b) | (c ^ d)) == 0)
(a != 0 && b != 0)  =>  (((a | -a) & (b | -b)) < 0)
(a != b && c != d)  =>  x = a ^ b; y = c ^ d;
                        (((x | -x) & (y | -y)) < 0)
(a < 0 || b < 0)    =>  ((a | b) < 0)
(a < 0 && b < 0)    =>  ((a & b) < 0)
(a >= 0 || b >= 0)  =>  ((a & b) >= 0)
```

```
(a >= 0 && b >= 0)  =>  ((a | b) >= 0)
(a < 0 || b >= 0)   =>  ((a | ~b) < 0)
(a < 0 && b >= 0)   =>  ((a & ~b) < 0)
(a >= 0 || b < 0)   =>  ((~a | b) < 0)
(a >= 0 && b < 0)   =>  ((~a & b) < 0)
(a != 0 && b < 0)   =>  (((a | -a) & b) < 0)
(a != 0 && b >= 0)  =>  (((a | -a) & ~b) < 0)
(a < 0 && b != 0)   =>  (((b | -b) & a) < 0)
(a >= 0 && b != 0)  =>  (((b | -b) & ~a) < 0)
```

`-fargument-alias`
`-fargument-noalias`
`-fargument-noalias-global`

Specify the possible relationships among parameters and between parameters and global data.

‘`-fargument-alias`’ specifies that arguments (parameters) may alias each other and may alias global storage. ‘`-fargument-noalias`’ specifies that arguments do not alias each other, but may alias global storage. ‘`-fargument-noalias-global`’ specifies that arguments do not alias each other and do not alias global storage.

Each language will automatically use whatever option is required by the language standard. You should not need to use these options yourself.

`-fleading-underscore`

This option and its counterpart, -fno-leading-underscore, forcibly change the way C symbols are represented in the object file. One use is to help link with legacy assembly code.

Be warned that you should know what you are doing when invoking this option, and that not all targets provide complete support for it.

2.17 Environment Variables Affecting GCC

This section describes several environment variables that affect how GCC operates. Some of them work by specifying directories or prefixes to use when searching for various kinds of files. Some are used to specify other aspects of the compilation environment.

Note that you can also specify places to search using options such as ‘`-B`’, ‘`-I`’ and ‘`-L`’ (see Section 2.12 [Directory Options], page 46). These take precedence over places specified using environment variables, which in turn take precedence over those specified by the configuration of GCC.

`LANG`
`LC_CTYPE`
`LC_MESSAGES`
`LC_ALL` These environment variables control the way that GCC uses localization information that allow GCC to work with different national conventions. GCC inspects the locale categories `LC_CTYPE` and `LC_MESSAGES` if it has been configured to do so. These locale categories can be set to any value supported by your installation. A typical value is ‘`en_UK`’ for English in the United Kingdom.

The `LC_CTYPE` environment variable specifies character classification. GCC uses it to determine the character boundaries in a string; this is needed for some multibyte encodings that contain quote and escape characters that would otherwise be interpreted as a string end or escape.

The `LC_MESSAGES` environment variable specifies the language to use in diagnostic messages.

If the `LC_ALL` environment variable is set, it overrides the value of `LC_CTYPE` and `LC_MESSAGES`; otherwise, `LC_CTYPE` and `LC_MESSAGES` default to the value of the `LANG` environment variable. If none of these variables are set, GCC defaults to traditional C English behavior.

`TMPDIR` If `TMPDIR` is set, it specifies the directory to use for temporary files. GCC uses temporary files to hold the output of one stage of compilation which is to be used as input to the next stage: for example, the output of the preprocessor, which is the input to the compiler proper.

`GCC_EXEC_PREFIX`

If `GCC_EXEC_PREFIX` is set, it specifies a prefix to use in the names of the subprograms executed by the compiler. No slash is added when this prefix is combined with the name of a subprogram, but you can specify a prefix that ends with a slash if you wish.

If `GCC_EXEC_PREFIX` is not set, GNU CC will attempt to figure out an appropriate prefix to use based on the pathname it was invoked with.

If GCC cannot find the subprogram using the specified prefix, it tries looking in the usual places for the subprogram.

The default value of `GCC_EXEC_PREFIX` is ‘*prefix*`/lib/gcc-lib/`’ where *prefix* is the value of `prefix` when you ran the ‘`configure`’ script.

Other prefixes specified with ‘`-B`’ take precedence over this prefix.

This prefix is also used for finding files such as ‘`crt0.o`’ that are used for linking.

In addition, the prefix is used in an unusual way in finding the directories to search for header files. For each of the standard directories whose name normally begins with ‘`/usr/local/lib/gcc-lib`’ (more precisely, with the value of `GCC_INCLUDE_DIR`), GCC tries replacing that beginning with the specified prefix to produce an alternate directory name. Thus, with ‘`-Bfoo/`’, GCC will search ‘`foo/bar`’ where it would normally search ‘`/usr/local/lib/bar`’. These alternate directories are searched first; the standard directories come next.

`COMPILER_PATH`

The value of `COMPILER_PATH` is a colon-separated list of directories, much like `PATH`. GCC tries the directories thus specified when searching for subprograms, if it can’t find the subprograms using `GCC_EXEC_PREFIX`.

`LIBRARY_PATH`

The value of `LIBRARY_PATH` is a colon-separated list of directories, much like `PATH`. When configured as a native compiler, GCC tries the directories thus specified when searching for special linker files, if it can’t find them using `GCC_EXEC_PREFIX`. Linking using GCC also uses these directories when searching for ordinary libraries for the ‘`-l`’ option (but directories specified with ‘`-L`’ come first).

`C_INCLUDE_PATH`
`CPLUS_INCLUDE_PATH`
`OBJC_INCLUDE_PATH`

These environment variables pertain to particular languages. Each variable's value is a colon-separated list of directories, much like `PATH`. When GCC searches for header files, it tries the directories listed in the variable for the language you are using, after the directories specified with '`-I`' but before the standard header file directories.

`DEPENDENCIES_OUTPUT`

If this variable is set, its value specifies how to output dependencies for Make based on the header files processed by the compiler. This output looks much like the output from the '`-M`' option (see Section 2.9 [Preprocessor Options], page 40), but it goes to a separate file, and is in addition to the usual results of compilation.

The value of `DEPENDENCIES_OUTPUT` can be just a file name, in which case the Make rules are written to that file, guessing the target name from the source file name. Or the value can have the form '*file target*', in which case the rules are written to file *file* using *target* as the target name.

`LANG`

This variable is used to pass locale information to the compiler. One way in which this information is used is to determine the character set to be used when character literals, string literals and comments are parsed in C and C++. When the compiler is configured to allow multibyte characters, the following values for `LANG` are recognized:

`C-JIS` Recognize JIS characters.

`C-SJIS` Recognize SJIS characters.

`C-EUCJP` Recognize EUCJP characters.

If `LANG` is not defined, or if it has some other value, then the compiler will use mblen and mbtowc as defined by the default locale to recognize and translate multibyte characters.

2.18 Running Protoize

The program `protoize` is an optional part of GNU C. You can use it to add prototypes to a program, thus converting the program to ANSI C in one respect. The companion program `unprotoize` does the reverse: it removes argument types from any prototypes that are found.

When you run these programs, you must specify a set of source files as command line arguments. The conversion programs start out by compiling these files to see what functions they define. The information gathered about a file *foo* is saved in a file named '*foo*`.X`'.

After scanning comes actual conversion. The specified files are all eligible to be converted; any files they include (whether sources or just headers) are eligible as well.

But not all the eligible files are converted. By default, `protoize` and `unprotoize` convert only source and header files in the current directory. You can specify additional directories whose files should be converted with the '`-d` *directory*' option. You can also specify particular files to exclude with the '`-x` *file*' option. A file is converted if it is eligible, its directory name matches one of the specified directory names, and its name within the directory has not been excluded.

Basic conversion with `protoize` consists of rewriting most function definitions and function declarations to specify the types of the arguments. The only ones not rewritten are those for varargs functions.

`protoize` optionally inserts prototype declarations at the beginning of the source file, to make them available for any calls that precede the function's definition. Or it can insert prototype declarations with block scope in the blocks where undeclared functions are called.

Basic conversion with `unprotoize` consists of rewriting most function declarations to remove any argument types, and rewriting function definitions to the old-style pre-ANSI form.

Both conversion programs print a warning for any function declaration or definition that they can't convert. You can suppress these warnings with '`-q`'.

The output from `protoize` or `unprotoize` replaces the original source file. The original file is renamed to a name ending with '`.save`' (for DOS, the saved filename ends in '`.sav`' without the original '`.c`' suffix). If the '`.save`' ('`.sav`' for DOS) file already exists, then the source file is simply discarded.

`protoize` and `unprotoize` both depend on GCC itself to scan the program and collect information about the functions it uses. So neither of these programs will work until GCC is installed.

Here is a table of the options you can use with `protoize` and `unprotoize`. Each option works with both programs unless otherwise stated.

`-B` *directory*
: Look for the file '`SYSCALLS.c.X`' in *directory*, instead of the usual directory (normally '`/usr/local/lib`'). This file contains prototype information about standard system functions. This option applies only to `protoize`.

`-c` *compilation-options*
: Use *compilation-options* as the options when running `gcc` to produce the '`.X`' files. The special option '`-aux-info`' is always passed in addition, to tell `gcc` to write a '`.X`' file.

  Note that the compilation options must be given as a single argument to `protoize` or `unprotoize`. If you want to specify several `gcc` options, you must quote the entire set of compilation options to make them a single word in the shell.

  There are certain `gcc` arguments that you cannot use, because they would produce the wrong kind of output. These include '`-g`', '`-O`', '`-c`', '`-S`', and '`-o`' If you include these in the *compilation-options*, they are ignored.

`-C`
: Rename files to end in '`.C`' ('`.cc`' for DOS-based file systems) instead of '`.c`'. This is convenient if you are converting a C program to C++. This option applies only to `protoize`.

`-g`
: Add explicit global declarations. This means inserting explicit declarations at the beginning of each source file for each function that is called in the file and was not declared. These declarations precede the first function definition that contains a call to an undeclared function. This option applies only to `protoize`.

`-i` *string*
: Indent old-style parameter declarations with the string *string*. This option applies only to `protoize`.

  `unprotoize` converts prototyped function definitions to old-style function definitions, where the arguments are declared between the argument list and the initial '`{`'. By

default, `unprotoize` uses five spaces as the indentation. If you want to indent with just one space instead, use '`-i " "`'.

`-k` Keep the '`.X`' files. Normally, they are deleted after conversion is finished.

`-l` Add explicit local declarations. `protoize` with '`-l`' inserts a prototype declaration for each function in each block which calls the function without any declaration. This option applies only to `protoize`.

`-n` Make no real changes. This mode just prints information about the conversions that would have been done without '`-n`'.

`-N` Make no '`.save`' files. The original files are simply deleted. Use this option with caution.

`-p` *program*
Use the program *program* as the compiler. Normally, the name '`gcc`' is used.

`-q` Work quietly. Most warnings are suppressed.

`-v` Print the version number, just like '`-v`' for `gcc`.

If you need special compiler options to compile one of your program's source files, then you should generate that file's '`.X`' file specially, by running `gcc` on that source file with the appropriate options and the option '`-aux-info`'. Then run `protoize` on the entire set of files. `protoize` will use the existing '`.X`' file because it is newer than the source file. For example:

```
gcc -Dfoo=bar file1.c -aux-info
protoize *.c
```

You need to include the special files along with the rest in the `protoize` command, even though their '`.X`' files already exist, because otherwise they won't get converted.

See Section 6.7 [Protoize Caveats], page 137, for more information on how to use `protoize` successfully.

3 Extensions to the C Language Family

GNU C provides several language features not found in ANSI standard C. (The '`-pedantic`' option directs GNU CC to print a warning message if any of these features is used.) To test for the availability of these features in conditional compilation, check for a predefined macro `__GNUC__`, which is always defined under GNU CC.

These extensions are available in C and Objective C. Most of them are also available in C++. See Chapter 4 [Extensions to the C++ Language], page 115, for extensions that apply *only* to C++.

3.1 Statements and Declarations in Expressions

A compound statement enclosed in parentheses may appear as an expression in GNU C. This allows you to use loops, switches, and local variables within an expression.

Recall that a compound statement is a sequence of statements surrounded by braces; in this construct, parentheses go around the braces. For example:

```
({ int y = foo (); int z;
   if (y > 0) z = y;
   else z = - y;
   z; })
```

is a valid (though slightly more complex than necessary) expression for the absolute value of `foo ()`.

The last thing in the compound statement should be an expression followed by a semicolon; the value of this subexpression serves as the value of the entire construct. (If you use some other kind of statement last within the braces, the construct has type `void`, and thus effectively no value.)

This feature is especially useful in making macro definitions "safe" (so that they evaluate each operand exactly once). For example, the "maximum" function is commonly defined as a macro in standard C as follows:

```
#define max(a,b) ((a) > (b) ? (a) : (b))
```

But this definition computes either *a* or *b* twice, with bad results if the operand has side effects. In GNU C, if you know the type of the operands (here let's assume `int`), you can define the macro safely as follows:

```
#define maxint(a,b) \
  ({int _a = (a), _b = (b); _a > _b ? _a : _b; })
```

Embedded statements are not allowed in constant expressions, such as the value of an enumeration constant, the width of a bit field, or the initial value of a static variable.

If you don't know the type of the operand, you can still do this, but you must use `typeof` (see Section 3.7 [Typeof], page 78) or type naming (see Section 3.6 [Naming Types], page 78).

3.2 Locally Declared Labels

Each statement expression is a scope in which *local labels* can be declared. A local label is simply an identifier; you can jump to it with an ordinary `goto` statement, but only from within the statement expression it belongs to.

A local label declaration looks like this:

```
__label__ label;
```

or

```
__label__ label1, label2, ...;
```

Local label declarations must come at the beginning of the statement expression, right after the '`({`', before any ordinary declarations.

The label declaration defines the label *name*, but does not define the label itself. You must do this in the usual way, with *label*`:`, within the statements of the statement expression.

The local label feature is useful because statement expressions are often used in macros. If the macro contains nested loops, a `goto` can be useful for breaking out of them. However, an ordinary label whose scope is the whole function cannot be used: if the macro can be expanded several times in one function, the label will be multiply defined in that function. A local label avoids this problem. For example:

```
#define SEARCH(array, target)                     \
({                                               \
  __label__ found;                                \
  typeof (target) _SEARCH_target = (target);      \
  typeof (*(array)) *_SEARCH_array = (array);     \
  int i, j;                                       \
  int value;                                      \
  for (i = 0; i < max; i++)                       \
    for (j = 0; j < max; j++)                     \
      if (_SEARCH_array[i][j] == _SEARCH_target)  \
        { value = i; goto found; }                \
  value = -1;                                     \
 found:                                           \
  value;                                          \
})
```

3.3 Labels as Values

You can get the address of a label defined in the current function (or a containing function) with the unary operator '`&&`'. The value has type `void *`. This value is a constant and can be used wherever a constant of that type is valid. For example:

```
void *ptr;
...
ptr = &&foo;
```

To use these values, you need to be able to jump to one. This is done with the computed goto statement[1], `goto *`*exp*`;`. For example,

```
goto *ptr;
```

Any expression of type `void *` is allowed.

One way of using these constants is in initializing a static array that will serve as a jump table:

```
static void *array[] = { &&foo, &&bar, &&hack };
```

Then you can select a label with indexing, like this:

[1] The analogous feature in Fortran is called an assigned goto, but that name seems inappropriate in C, where one can do more than simply store label addresses in label variables.

```
goto *array[i];
```

Note that this does not check whether the subscript is in bounds—array indexing in C never does that.

Such an array of label values serves a purpose much like that of the `switch` statement. The `switch` statement is cleaner, so use that rather than an array unless the problem does not fit a `switch` statement very well.

Another use of label values is in an interpreter for threaded code. The labels within the interpreter function can be stored in the threaded code for super-fast dispatching.

You may not use this mechanism to jump to code in a different function. If you do that, totally unpredictable things will happen. The best way to avoid this is to store the label address only in automatic variables and never pass it as an argument.

An alternate way to write the above example is

```
static const int array[] = { &&foo - &&foo, &&bar - &&foo, &&hack - &&foo };
goto *(&&foo + array[i]);
```

This is more friendly to code living in shared libraries, as it reduces the number of dynamic relocations that are needed, and by consequence, allows the data to be read-only.

3.4 Nested Functions

A *nested function* is a function defined inside another function. (Nested functions are not supported for GNU C++.) The nested function's name is local to the block where it is defined. For example, here we define a nested function named `square`, and call it twice:

```
foo (double a, double b)
{
  double square (double z) { return z * z; }

  return square (a) + square (b);
}
```

The nested function can access all the variables of the containing function that are visible at the point of its definition. This is called *lexical scoping*. For example, here we show a nested function which uses an inherited variable named `offset`:

```
bar (int *array, int offset, int size)
{
  int access (int *array, int index)
    { return array[index + offset]; }
  int i;
  ...
  for (i = 0; i < size; i++)
    ... access (array, i) ...
}
```

Nested function definitions are permitted within functions in the places where variable definitions are allowed; that is, in any block, before the first statement in the block.

It is possible to call the nested function from outside the scope of its name by storing its address or passing the address to another function:

```
hack (int *array, int size)
{
  void store (int index, int value)
    { array[index] = value; }

  intermediate (store, size);
}
```

Here, the function `intermediate` receives the address of `store` as an argument. If `intermediate` calls `store`, the arguments given to `store` are used to store into `array`. But this technique works only so long as the containing function (`hack`, in this example) does not exit.

If you try to call the nested function through its address after the containing function has exited, all hell will break loose. If you try to call it after a containing scope level has exited, and if it refers to some of the variables that are no longer in scope, you may be lucky, but it's not wise to take the risk. If, however, the nested function does not refer to anything that has gone out of scope, you should be safe.

GNU CC implements taking the address of a nested function using a technique called *trampolines*. A paper describing them is available as '`http://master.debian.org/~karlheg/Usenix88-lexic.pdf`'.

A nested function can jump to a label inherited from a containing function, provided the label was explicitly declared in the containing function (see Section 3.2 [Local Labels], page 74). Such a jump returns instantly to the containing function, exiting the nested function which did the `goto` and any intermediate functions as well. Here is an example:

```
bar (int *array, int offset, int size)
{
  __label__ failure;
  int access (int *array, int index)
    {
      if (index > size)
        goto failure;
      return array[index + offset];
    }
  int i;
  ...
  for (i = 0; i < size; i++)
    ... access (array, i) ...
  ...
  return 0;

 /* Control comes here from access
  if it detects an error.  */
 failure:
  return -1;
}
```

A nested function always has internal linkage. Declaring one with `extern` is erroneous. If you need to declare the nested function before its definition, use `auto` (which is otherwise meaningless for function declarations).

```
bar (int *array, int offset, int size)
{
```

```
    __label__ failure;
    auto int access (int *, int);
    ...
    int access (int *array, int index)
      {
        if (index > size)
          goto failure;
        return array[index + offset];
      }
    ...
  }
```

3.5 Constructing Function Calls

Using the built-in functions described below, you can record the arguments a function received, and call another function with the same arguments, without knowing the number or types of the arguments.

You can also record the return value of that function call, and later return that value, without knowing what data type the function tried to return (as long as your caller expects that data type).

`__builtin_apply_args ()`
: This built-in function returns a pointer of type `void *` to data describing how to perform a call with the same arguments as were passed to the current function.

  The function saves the arg pointer register, structure value address, and all registers that might be used to pass arguments to a function into a block of memory allocated on the stack. Then it returns the address of that block.

`__builtin_apply (`*function*`,` *arguments*`,` *size*`)`
: This built-in function invokes *function* (type `void (*)()`) with a copy of the parameters described by *arguments* (type `void *`) and *size* (type `int`).

  The value of *arguments* should be the value returned by `__builtin_apply_args`. The argument *size* specifies the size of the stack argument data, in bytes.

  This function returns a pointer of type `void *` to data describing how to return whatever value was returned by *function*. The data is saved in a block of memory allocated on the stack.

  It is not always simple to compute the proper value for *size*. The value is used by `__builtin_apply` to compute the amount of data that should be pushed on the stack and copied from the incoming argument area.

`__builtin_return (`*result*`)`
: This built-in function returns the value described by *result* from the containing function. You should specify, for *result*, a value returned by `__builtin_apply`.

3.6 Naming an Expression's Type

You can give a name to the type of an expression using a `typedef` declaration with an initializer. Here is how to define *name* as a type name for the type of *exp*:

```
typedef name = exp;
```

This is useful in conjunction with the statements-within-expressions feature. Here is how the two together can be used to define a safe "maximum" macro that operates on any arithmetic type:

```
#define max(a,b) \
  ({typedef _ta = (a), _tb = (b);  \
    _ta _a = (a); _tb _b = (b);     \
    _a > _b ? _a : _b; })
```

The reason for using names that start with underscores for the local variables is to avoid conflicts with variable names that occur within the expressions that are substituted for `a` and `b`. Eventually we hope to design a new form of declaration syntax that allows you to declare variables whose scopes start only after their initializers; this will be a more reliable way to prevent such conflicts.

3.7 Referring to a Type with `typeof`

Another way to refer to the type of an expression is with `typeof`. The syntax of using of this keyword looks like `sizeof`, but the construct acts semantically like a type name defined with `typedef`.

There are two ways of writing the argument to `typeof`: with an expression or with a type. Here is an example with an expression:

```
typeof (x[0](1))
```

This assumes that `x` is an array of functions; the type described is that of the values of the functions.

Here is an example with a typename as the argument:

```
typeof (int *)
```

Here the type described is that of pointers to `int`.

If you are writing a header file that must work when included in ANSI C programs, write `__typeof__` instead of `typeof`. See Section 3.36 [Alternate Keywords], page 109.

A `typeof`-construct can be used anywhere a typedef name could be used. For example, you can use it in a declaration, in a cast, or inside of `sizeof` or `typeof`.

- This declares `y` with the type of what `x` points to.

  ```
  typeof (*x) y;
  ```

- This declares `y` as an array of such values.

  ```
  typeof (*x) y[4];
  ```

- This declares `y` as an array of pointers to characters:

  ```
  typeof (typeof (char *)[4]) y;
  ```

  It is equivalent to the following traditional C declaration:

  ```
  char *y[4];
  ```

  To see the meaning of the declaration using `typeof`, and why it might be a useful way to write, let's rewrite it with these macros:

  ```
  #define pointer(T)  typeof(T *)
  #define array(T, N) typeof(T [N])
  ```

  Now the declaration can be rewritten this way:

  ```
  array (pointer (char), 4) y;
  ```

  Thus, `array (pointer (char), 4)` is the type of arrays of 4 pointers to `char`.

3.8 Generalized Lvalues

Compound expressions, conditional expressions and casts are allowed as lvalues provided their operands are lvalues. This means that you can take their addresses or store values into them.

Standard C++ allows compound expressions and conditional expressions as lvalues, and permits casts to reference type, so use of this extension is deprecated for C++ code.

For example, a compound expression can be assigned, provided the last expression in the sequence is an lvalue. These two expressions are equivalent:

```
(a, b) += 5
a, (b += 5)
```

Similarly, the address of the compound expression can be taken. These two expressions are equivalent:

```
&(a, b)
a, &b
```

A conditional expression is a valid lvalue if its type is not void and the true and false branches are both valid lvalues. For example, these two expressions are equivalent:

```
(a ? b : c) = 5
(a ? b = 5 : (c = 5))
```

A cast is a valid lvalue if its operand is an lvalue. A simple assignment whose left-hand side is a cast works by converting the right-hand side first to the specified type, then to the type of the inner left-hand side expression. After this is stored, the value is converted back to the specified type to become the value of the assignment. Thus, if a has type `char *`, the following two expressions are equivalent:

```
(int)a = 5
(int)(a = (char *)(int)5)
```

An assignment-with-arithmetic operation such as '`+=`' applied to a cast performs the arithmetic using the type resulting from the cast, and then continues as in the previous case. Therefore, these two expressions are equivalent:

```
(int)a += 5
(int)(a = (char *)(int) ((int)a + 5))
```

You cannot take the address of an lvalue cast, because the use of its address would not work out coherently. Suppose that `&(int)f` were permitted, where `f` has type `float`. Then the following statement would try to store an integer bit-pattern where a floating point number belongs:

```
*&(int)f = 1;
```

This is quite different from what `(int)f = 1` would do—that would convert 1 to floating point and store it. Rather than cause this inconsistency, we think it is better to prohibit use of '`&`' on a cast.

If you really do want an `int *` pointer with the address of `f`, you can simply write `(int *)&f`.

3.9 Conditionals with Omitted Operands

The middle operand in a conditional expression may be omitted. Then if the first operand is nonzero, its value is the value of the conditional expression.

Therefore, the expression

```
x ? : y
```

has the value of `x` if that is nonzero; otherwise, the value of `y`.

This example is perfectly equivalent to

```
x ? x : y
```

In this simple case, the ability to omit the middle operand is not especially useful. When it becomes useful is when the first operand does, or may (if it is a macro argument), contain a side effect. Then repeating the operand in the middle would perform the side effect twice. Omitting the middle operand uses the value already computed without the undesirable effects of recomputing it.

3.10 Double-Word Integers

GNU C supports data types for integers that are twice as long as `int`. Simply write `long long int` for a signed integer, or `unsigned long long int` for an unsigned integer. To make an integer constant of type `long long int`, add the suffix `LL` to the integer. To make an integer constant of type `unsigned long long int`, add the suffix `ULL` to the integer.

You can use these types in arithmetic like any other integer types. Addition, subtraction, and bitwise boolean operations on these types are open-coded on all types of machines. Multiplication is open-coded if the machine supports fullword-to-doubleword a widening multiply instruction. Division and shifts are open-coded only on machines that provide special support. The operations that are not open-coded use special library routines that come with GNU CC.

There may be pitfalls when you use `long long` types for function arguments, unless you declare function prototypes. If a function expects type `int` for its argument, and you pass a value of type `long long int`, confusion will result because the caller and the subroutine will disagree about the number of bytes for the argument. Likewise, if the function expects `long long int` and you pass `int`. The best way to avoid such problems is to use prototypes.

3.11 Complex Numbers

GNU C supports complex data types. You can declare both complex integer types and complex floating types, using the keyword `__complex__`.

For example, ‘`__complex__ double x;`’ declares `x` as a variable whose real part and imaginary part are both of type `double`. ‘`__complex__ short int y;`’ declares `y` to have real and imaginary parts of type `short int`; this is not likely to be useful, but it shows that the set of complex types is complete.

To write a constant with a complex data type, use the suffix ‘`i`’ or ‘`j`’ (either one; they are equivalent). For example, `2.5fi` has type `__complex__ float` and `3i` has type `__complex__ int`. Such a constant always has a pure imaginary value, but you can form any complex value you like by adding one to a real constant.

To extract the real part of a complex-valued expression *exp*, write `__real__` *exp*. Likewise, use `__imag__` to extract the imaginary part.

The operator ‘`~`’ performs complex conjugation when used on a value with a complex type.

GNU CC can allocate complex automatic variables in a noncontiguous fashion; it’s even possible for the real part to be in a register while the imaginary part is on the stack (or vice-versa). None of the supported debugging info formats has a way to represent noncontiguous allocation like this, so GNU

CC describes a noncontiguous complex variable as if it were two separate variables of noncomplex type. If the variable's actual name is `foo`, the two fictitious variables are named `foo$real` and `foo$imag`. You can examine and set these two fictitious variables with your debugger.

A future version of GDB will know how to recognize such pairs and treat them as a single variable with a complex type.

3.12 Hex Floats

GNU CC recognizes floating-point numbers writen not only in the usual decimal notation, such as `1.55e1`, but also numbers such as `0x1.fp3` written in hexadecimal format. In that format the `0x` hex introducer and the `p` or `P` exponent field are mandatory. The exponent is a decimal number that indicates the power of 2 by which the significand part will be multiplied. Thus `0x1.f` is 1 15/16, `p3` multiplies it by 8, and the value of `0x1.fp3` is the same as `1.55e1`.

Unlike for floating-point numbers in the decimal notation the exponent is always required in the hexadecimal notation. Otherwise the compiler would not be able to resolve the ambiguity of, e.g., `0x1.f`. This could mean `1.0f` or `1.9375` since `f` is also the extension for floating-point constants of type `float`.

3.13 Arrays of Length Zero

Zero-length arrays are allowed in GNU C. They are very useful as the last element of a structure which is really a header for a variable-length object:

```
struct line {
  int length;
  char contents[0];
};

{
  struct line *thisline = (struct line *)
    malloc (sizeof (struct line) + this_length);
  thisline->length = this_length;
}
```

In standard C, you would have to give `contents` a length of 1, which means either you waste space or complicate the argument to `malloc`.

3.14 Arrays of Variable Length

Variable-length automatic arrays are allowed in GNU C. These arrays are declared like any other automatic arrays, but with a length that is not a constant expression. The storage is allocated at the point of declaration and deallocated when the brace-level is exited. For example:

```
FILE *
concat_fopen (char *s1, char *s2, char *mode)
{
  char str[strlen (s1) + strlen (s2) + 1];
  strcpy (str, s1);
  strcat (str, s2);
```

```
  return fopen (str, mode);
}
```

Jumping or breaking out of the scope of the array name deallocates the storage. Jumping into the scope is not allowed; you get an error message for it.

You can use the function `alloca` to get an effect much like variable-length arrays. The function `alloca` is available in many other C implementations (but not in all). On the other hand, variable-length arrays are more elegant.

There are other differences between these two methods. Space allocated with `alloca` exists until the containing *function* returns. The space for a variable-length array is deallocated as soon as the array name's scope ends. (If you use both variable-length arrays and `alloca` in the same function, deallocation of a variable-length array will also deallocate anything more recently allocated with `alloca`.)

You can also use variable-length arrays as arguments to functions:

```
struct entry
tester (int len, char data[len][len])
{
  ...
}
```

The length of an array is computed once when the storage is allocated and is remembered for the scope of the array in case you access it with `sizeof`.

If you want to pass the array first and the length afterward, you can use a forward declaration in the parameter list—another GNU extension.

```
struct entry
tester (int len; char data[len][len], int len)
{
  ...
}
```

The '`int len`' before the semicolon is a *parameter forward declaration*, and it serves the purpose of making the name `len` known when the declaration of `data` is parsed.

You can write any number of such parameter forward declarations in the parameter list. They can be separated by commas or semicolons, but the last one must end with a semicolon, which is followed by the "real" parameter declarations. Each forward declaration must match a "real" declaration in parameter name and data type.

3.15 Macros with Variable Numbers of Arguments

In GNU C, a macro can accept a variable number of arguments, much as a function can. The syntax for defining the macro looks much like that used for a function. Here is an example:

```
#define eprintf(format, args...)  \
 fprintf (stderr, format , ## args)
```

Here `args` is a *rest argument*: it takes in zero or more arguments, as many as the call contains. All of them plus the commas between them form the value of `args`, which is substituted into the macro body where `args` is used. Thus, we have this expansion:

```
eprintf ("%s:%d: ", input_file_name, line_number)
↦
fprintf (stderr, "%s:%d: " , input_file_name, line_number)
```

Note that the comma after the string constant comes from the definition of `eprintf`, whereas the last comma comes from the value of `args`.

The reason for using '`##`' is to handle the case when `args` matches no arguments at all. In this case, `args` has an empty value. In this case, the second comma in the definition becomes an embarrassment: if it got through to the expansion of the macro, we would get something like this:

```
fprintf (stderr, "success!\n" , )
```

which is invalid C syntax. '`##`' gets rid of the comma, so we get the following instead:

```
fprintf (stderr, "success!\n")
```

This is a special feature of the GNU C preprocessor: '`##`' before a rest argument that is empty discards the preceding sequence of non-whitespace characters from the macro definition. (If another macro argument precedes, none of it is discarded.)

It might be better to discard the last preprocessor token instead of the last preceding sequence of non-whitespace characters; in fact, we may someday change this feature to do so. We advise you to write the macro definition so that the preceding sequence of non-whitespace characters is just a single token, so that the meaning will not change if we change the definition of this feature.

3.16 Non-Lvalue Arrays May Have Subscripts

Subscripting is allowed on arrays that are not lvalues, even though the unary '`&`' operator is not. For example, this is valid in GNU C though not valid in other C dialects:

```
struct foo {int a[4];};

struct foo f();

bar (int index)
{
  return f().a[index];
}
```

3.17 Arithmetic on `void`- and Function-Pointers

In GNU C, addition and subtraction operations are supported on pointers to `void` and on pointers to functions. This is done by treating the size of a `void` or of a function as 1.

A consequence of this is that `sizeof` is also allowed on `void` and on function types, and returns 1.

The option '`-Wpointer-arith`' requests a warning if these extensions are used.

3.18 Non-Constant Initializers

As in standard C++, the elements of an aggregate initializer for an automatic variable are not required to be constant expressions in GNU C. Here is an example of an initializer with run-time varying elements:

```
foo (float f, float g)
{
  float beat_freqs[2] = { f-g, f+g };
  ...
}
```

3.19 Constructor Expressions

GNU C supports constructor expressions. A constructor looks like a cast containing an initializer. Its value is an object of the type specified in the cast, containing the elements specified in the initializer.

Usually, the specified type is a structure. Assume that `struct foo` and `structure` are declared as shown:

```
struct foo {int a; char b[2];} structure;
```

Here is an example of constructing a `struct foo` with a constructor:

```
structure = ((struct foo) {x + y, 'a', 0});
```

This is equivalent to writing the following:

```
{
  struct foo temp = {x + y, 'a', 0};
  structure = temp;
}
```

You can also construct an array. If all the elements of the constructor are (made up of) simple constant expressions, suitable for use in initializers, then the constructor is an lvalue and can be coerced to a pointer to its first element, as shown here:

```
char **foo = (char *[]) { "x", "y", "z" };
```

Array constructors whose elements are not simple constants are not very useful, because the constructor is not an lvalue. There are only two valid ways to use it: to subscript it, or initialize an array variable with it. The former is probably slower than a `switch` statement, while the latter does the same thing an ordinary C initializer would do. Here is an example of subscripting an array constructor:

```
output = ((int[]) { 2, x, 28 }) [input];
```

Constructor expressions for scalar types and union types are is also allowed, but then the constructor expression is equivalent to a cast.

3.20 Labeled Elements in Initializers

Standard C requires the elements of an initializer to appear in a fixed order, the same as the order of the elements in the array or structure being initialized.

In GNU C you can give the elements in any order, specifying the array indices or structure field names they apply to. This extension is not implemented in GNU C++.

To specify an array index, write '`[`*index*`]`' or '`[`*index*`] =`' before the element value. For example,

```
int a[6] = { [4] 29, [2] = 15 };
```

is equivalent to

```
int a[6] = { 0, 0, 15, 0, 29, 0 };
```

The index values must be constant expressions, even if the array being initialized is automatic.

To initialize a range of elements to the same value, write '`[`*first* `...` *last*`] =` *value*'. For example,

```
int widths[] = { [0 ... 9] = 1, [10 ... 99] = 2, [100] = 3 };
```

Note that the length of the array is the highest value specified plus one.

In a structure initializer, specify the name of a field to initialize with '*fieldname*`:`' before the element value. For example, given the following structure,

```
struct point { int x, y; };
```

the following initialization

```
struct point p = { y: yvalue, x: xvalue };
```

is equivalent to

```
struct point p = { xvalue, yvalue };
```

Another syntax which has the same meaning is '`.`*fieldname* `=`'., as shown here:

```
struct point p = { .y = yvalue, .x = xvalue };
```

You can also use an element label (with either the colon syntax or the period-equal syntax) when initializing a union, to specify which element of the union should be used. For example,

```
union foo { int i; double d; };

union foo f = { d: 4 };
```

will convert 4 to a `double` to store it in the union using the second element. By contrast, casting 4 to type `union foo` would store it into the union as the integer `i`, since it is an integer. (See Section 3.22 [Cast to Union], page 86.)

You can combine this technique of naming elements with ordinary C initialization of successive elements. Each initializer element that does not have a label applies to the next consecutive element of the array or structure. For example,

```
int a[6] = { [1] = v1, v2, [4] = v4 };
```

is equivalent to

```
int a[6] = { 0, v1, v2, 0, v4, 0 };
```

Labeling the elements of an array initializer is especially useful when the indices are characters or belong to an `enum` type. For example:

```
int whitespace[256]
  = { [' '] = 1, ['\t'] = 1, ['\h'] = 1,
      ['\f'] = 1, ['\n'] = 1, ['\r'] = 1 };
```

3.21 Case Ranges

You can specify a range of consecutive values in a single `case` label, like this:

`case` *low* `...` *high*`:`

This has the same effect as the proper number of individual `case` labels, one for each integer value from *low* to *high*, inclusive.

This feature is especially useful for ranges of ASCII character codes:

```
case 'A' ... 'Z':
```

Be careful: Write spaces around the ..., for otherwise it may be parsed wrong when you use it with integer values. For example, write this:

```
case 1 ... 5:
```

rather than this:

```
case 1...5:
```

3.22 Cast to a Union Type

A cast to union type is similar to other casts, except that the type specified is a union type. You can specify the type either with `union` *tag* or with a typedef name. A cast to union is actually a constructor though, not a cast, and hence does not yield an lvalue like normal casts. (See Section 3.19 [Constructors], page 84.)

The types that may be cast to the union type are those of the members of the union. Thus, given the following union and variables:

```
union foo { int i; double d; };
int x;
double y;
```

both `x` and `y` can be cast to type `union` foo.

Using the cast as the right-hand side of an assignment to a variable of union type is equivalent to storing in a member of the union:

```
union foo u;
...
u = (union foo) x  ≡  u.i = x
u = (union foo) y  ≡  u.d = y
```

You can also use the union cast as a function argument:

```
void hack (union foo);
...
hack ((union foo) x);
```

3.23 Declaring Attributes of Functions

In GNU C, you declare certain things about functions called in your program which help the compiler optimize function calls and check your code more carefully.

The keyword `__attribute__` allows you to specify special attributes when making a declaration. This keyword is followed by an attribute specification inside double parentheses. Ten attributes, `noreturn`, `const`, `format`, `no_instrument_function`, `section`, `constructor`, `destructor`, `unused`, `weak` and `malloc` are currently defined for functions. Other attributes, including `section` are supported for variables declarations (see Section 3.29 [Variable Attributes], page 92) and for types (see Section 3.30 [Type Attributes], page 95).

You may also specify attributes with '__' preceding and following each keyword. This allows you to use them in header files without being concerned about a possible macro of the same name. For example, you may use `__noreturn__` instead of `noreturn`.

`noreturn` A few standard library functions, such as `abort` and `exit`, cannot return. GNU CC knows this automatically. Some programs define their own functions that never return. You can declare them `noreturn` to tell the compiler this fact. For example,

```
void fatal () __attribute__ ((noreturn));

void
fatal (...)
{
  ... /* Print error message. */ ...
  exit (1);
}
```

The `noreturn` keyword tells the compiler to assume that `fatal` cannot return. It can then optimize without regard to what would happen if `fatal` ever did return. This makes slightly better code. More importantly, it helps avoid spurious warnings of uninitialized variables.

Do not assume that registers saved by the calling function are restored before calling the `noreturn` function.

It does not make sense for a `noreturn` function to have a return type other than `void`.

The attribute `noreturn` is not implemented in GNU C versions earlier than 2.5. An alternative way to declare that a function does not return, which works in the current version and in some older versions, is as follows:

```
typedef void voidfn ();

volatile voidfn fatal;
```

`pure` Many functions have no effects except the return value and their return value depends only on the parameters and/or global variables. Such a function can be subject to common subexpression elimination and loop optimization just as an arithmetic operator would be. These functions should be declared with the attribute `pure`. For example,

```
int square (int) __attribute__ ((pure));
```

says that the hypothetical function `square` is safe to call fewer times than the program says.

Some of common examples of pure functions are `strlen` or `memcmp`. Interesting non-pure functions are functions with infinite loops or those depending on volatile memory or other system resource, that may change between two consetuctive calls (such as `feof` in multithreding environment).

The attribute `pure` is not implemented in GNU C versions earlier than 2.96.

`const` Many functions do not examine any values except their arguments, and have no effects except the return value. Basically this is just slightly more strict class than the "pure" attribute above, since function is not alloved to read global memory.

Note that a function that has pointer arguments and examines the data pointed to must *not* be declared `const`. Likewise, a function that calls a non-`const` function usually must not be `const`. It does not make sense for a `const` function to return `void`.

The attribute `const` is not implemented in GNU C versions earlier than 2.5. An alternative way to declare that a function has no side effects, which works in the current version and in some older versions, is as follows:

```
typedef int intfn ();

extern const intfn square;
```

This approach does not work in GNU C++ from 2.6.0 on, since the language specifies that the '`const`' must be attached to the return value.

`format` (*archetype*, *string-index*, *first-to-check*)

The `format` attribute specifies that a function takes `printf`, `scanf`, or `strftime` style arguments which should be type-checked against a format string. For example, the declaration:

```
extern int
my_printf (void *my_object, const char *my_format, ...)
      __attribute__ ((format (printf, 2, 3)));
```

causes the compiler to check the arguments in calls to `my_printf` for consistency with the `printf` style format string argument `my_format`.

The parameter *archetype* determines how the format string is interpreted, and should be either `printf`, `scanf`, or `strftime`. The parameter *string-index* specifies which argument is the format string argument (starting from 1), while *first-to-check* is the number of the first argument to check against the format string. For functions where the arguments are not available to be checked (such as `vprintf`), specify the third parameter as zero. In this case the compiler only checks the format string for consistency.

In the example above, the format string (`my_format`) is the second argument of the function `my_print`, and the arguments to check start with the third argument, so the correct parameters for the format attribute are 2 and 3.

The `format` attribute allows you to identify your own functions which take format strings as arguments, so that GNU CC can check the calls to these functions for errors. The compiler always checks formats for the ANSI library functions `printf`, `fprintf`, `sprintf`, `scanf`, `fscanf`, `sscanf`, `strftime`, `vprintf`, `vfprintf` and `vsprintf` whenever such warnings are requested (using '`-Wformat`'), so there is no need to modify the header file '`stdio.h`'.

`format_arg` (*string-index*)

The `format_arg` attribute specifies that a function takes `printf` or `scanf` style arguments, modifies it (for example, to translate it into another language), and passes it to a `printf` or `scanf` style function. For example, the declaration:

```
extern char *
my_dgettext (char *my_domain, const char *my_format)
      __attribute__ ((format_arg (2)));
```

causes the compiler to check the arguments in calls to `my_dgettext` whose result is passed to a `printf`, `scanf`, or `strftime` type function for consistency with the `printf` style format string argument `my_format`.

The parameter *string-index* specifies which argument is the format string argument (starting from 1).

The `format-arg` attribute allows you to identify your own functions which modify format strings, so that GNU CC can check the calls to `printf`, `scanf`, or `strftime` function whose operands are a call to one of your own function. The compiler always treats `gettext`, `dgettext`, and `dcgettext` in this manner.

`no_instrument_function`

If '`-finstrument-functions`' is given, profiling function calls will be generated at entry and exit of most user-compiled functions. Functions with this attribute will not be so instrumented.

`section ("section-name")`

Normally, the compiler places the code it generates in the `text` section. Sometimes, however, you need additional sections, or you need certain particular functions to appear in special sections. The `section` attribute specifies that a function lives in a particular section. For example, the declaration:

```
extern void foobar (void) __attribute__ ((section ("bar")));
```

puts the function `foobar` in the `bar` section.

The ELF object file format permits arbitrary sections, so you can use this feature on MIPS SDE. However, you'll need to write a special linker script to do anything useful with the many sections generated.

If you need to map the entire contents of a module to a particular section, the linker can do that. And if you wanted to put every function in its own section for total link-time control, you should have used the option '`-function-sections`'.

`constructor`
`destructor`

The `constructor` attribute causes the function to be called automatically before execution enters `main ()`. Similarly, the `destructor` attribute causes the function to be called automatically after `main ()` has completed or `exit ()` has been called. Functions with these attributes are useful for initializing data that will be used implicitly during the execution of the program.

These attributes are not currently implemented for Objective C.

`unused` This attribute, attached to a function, means that the function is meant to be possibly unused. GNU CC will not produce a warning for this function. GNU C++ does not currently support this attribute as definitions without parameters are valid in C++.

`weak` The `weak` attribute causes the declaration to be emitted as a weak symbol rather than a global. This is primarily useful in defining library functions which can be overridden in user code, though it can also be used with non-function declarations. Weak symbols are supported for ELF targets, and also for a.out targets when using the GNU assembler and linker.

`malloc` The `malloc` attribute is used to tell the compiler that a function may be treated as if it were the malloc function. The compiler assumes that calls to malloc result in a pointers that cannot alias anything. This will often improve optimization.

`alias ("target")`

The `alias` attribute causes the declaration to be emitted as an alias for another symbol, which must be specified. For instance,

```
void __f () { /* do something */; }
void f () __attribute__ ((weak, alias ("__f")));
```

declares '`f`' to be a weak alias for '`__f`'. In C++, the mangled name for the target must be used.

Not all target machines support this attribute.

`no_check_memory_usage`
: The `no_check_memory_usage` attribute causes GNU CC to omit checks of memory references when it generates code for that function. Normally if you specify '`-fcheck-memory-usage`' (see see Section 2.16 [Code Gen Options], page 61), GNU CC generates calls to support routines before most memory accesses to permit support code to record usage and detect uses of uninitialized or unallocated storage. Since GNU CC cannot handle `asm` statements properly they are not allowed in such functions. If you declare a function with this attribute, GNU CC will not generate memory checking code for that function, permitting the use of `asm` statements without having to compile that function with different options. This also allows you to write support routines of your own if you wish, without getting infinite recursion if they get compiled with `-fcheck-memory-usage`.

`longcall`
: On MIPS CPUs, the `longcall` attribute causes the compiler to always call the function via a pointer, so that functions which don't reside in the same 256Mbyte aligned memory region as the current location can be called.

`mips16`
: On MIPS CPUs, the `mips16` attribute causes the compiler to generate code for this function using the compressed 16-bit MIPS16 ISA. If the base ISA is MIPS32 or MIPS64 then this function will generate MIPS16e code, for other ISAs MIPS16.

`nomips16`
: On MIPS CPUs, the `nomips16` attribute causes the compiler to generate code for this function using the base 32-bit MIPS ISA. If no 32-bit ISA was selected on the compiler command line, only MIPS16 or MIPS16e, then this function will generate MIPS I code in the former case, and MIPS32 in the latter.

You can specify multiple attributes in a declaration by separating them by commas within the double parentheses or by immediately following an attribute declaration with another attribute declaration.

Some people object to the `__attribute__` feature, suggesting that ANSI C's `#pragma` should be used instead. There are two reasons for not doing this.

1. It is impossible to generate `#pragma` commands from a macro.
2. There is no telling what the same `#pragma` might mean in another compiler.

These two reasons apply to almost any application that might be proposed for `#pragma`. It is basically a mistake to use `#pragma` for *anything*.

3.24 Prototypes and Old-Style Function Definitions

GNU C extends ANSI C to allow a function prototype to override a later old-style non-prototype definition. Consider the following example:

```
/* Use prototypes unless the compiler is old-fashioned.  */
#ifdef __STDC__
#define P(x) x
#else
#define P(x) ()
#endif
```

```
/* Prototype function declaration.  */
int isroot P((uid_t));

/* Old-style function definition.  */
int
isroot (x)   /* ??? lossage here ??? */
     uid_t x;
{
  return x == 0;
}
```

Suppose the type `uid_t` happens to be `short`. ANSI C does not allow this example, because subword arguments in old-style non-prototype definitions are promoted. Therefore in this example the function definition's argument is really an `int`, which does not match the prototype argument type of `short`.

This restriction of ANSI C makes it hard to write code that is portable to traditional C compilers, because the programmer does not know whether the `uid_t` type is `short`, `int`, or `long`. Therefore, in cases like these GNU C allows a prototype to override a later old-style definition. More precisely, in GNU C, a function prototype argument type overrides the argument type specified by a later old-style definition if the former type is the same as the latter type before promotion. Thus in GNU C the above example is equivalent to the following:

```
int isroot (uid_t);

int
isroot (uid_t x)
{
  return x == 0;
}
```

GNU C++ does not support old-style function definitions, so this extension is irrelevant.

3.25 C++ Style Comments

In GNU C, you may use C++ style comments, which start with '`//`' and continue until the end of the line. Many other C implementations allow such comments, and they are likely to be in a future C standard. However, C++ style comments are not recognized if you specify '`-ansi`' or '`-traditional`', since they are incompatible with traditional constructs like `dividend//*comment*/divisor`.

3.26 Dollar Signs in Identifier Names

In GNU C, you may normally use dollar signs in identifier names. This is because many traditional C implementations allow such identifiers. However, dollar signs in identifiers are not supported on a few target machines, typically because the target assembler does not allow them.

3.27 The Character ⟨ESC⟩ in Constants

You can use the sequence '`\e`' in a string or character constant to stand for the ASCII character ⟨ESC⟩.

3.28 Inquiring on Alignment of Types or Variables

The keyword `__alignof__` allows you to inquire about how an object is aligned, or the minimum alignment usually required by a type. Its syntax is just like `sizeof`.

For example, if the target machine requires a `double` value to be aligned on an 8-byte boundary, then `__alignof__ (double)` is 8. This is true on many RISC machines. On more traditional machine designs, `__alignof__ (double)` is 4 or even 2.

Some machines never actually require alignment; they allow reference to any data type even at an odd addresses. For these machines, `__alignof__` reports the *recommended* alignment of a type.

When the operand of `__alignof__` is an lvalue rather than a type, the value is the largest alignment that the lvalue is known to have. It may have this alignment as a result of its data type, or because it is part of a structure and inherits alignment from that structure. For example, after this declaration:

```
struct foo { int x; char y; } foo1;
```

the value of `__alignof__ (foo1.y)` is probably 2 or 4, the same as `__alignof__ (int)`, even though the data type of `foo1.y` does not itself demand any alignment.

It is an error to ask for the alignment of an incomplete type.

A related feature which lets you specify the alignment of an object is `__attribute__ ((aligned (`*alignment*`)))`; see the following section.

3.29 Specifying Attributes of Variables

The keyword `__attribute__` allows you to specify special attributes of variables or structure fields. This keyword is followed by an attribute specification inside double parentheses. Eight attributes are currently defined for variables: `aligned`, `mode`, `nocommon`, `packed`, `section`, `transparent_union`, `unused`, and `weak`. Other attributes are available for functions (see Section 3.23 [Function Attributes], page 86) and for types (see Section 3.30 [Type Attributes], page 95).

You may also specify attributes with '`__`' preceding and following each keyword. This allows you to use them in header files without being concerned about a possible macro of the same name. For example, you may use `__aligned__` instead of `aligned`.

`aligned (`*alignment*`)`

This attribute specifies a minimum alignment for the variable or structure field, measured in bytes. For example, the declaration:

```
int x __attribute__ ((aligned (16))) = 0;
```

causes the compiler to allocate the global variable `x` on a 16-byte boundary. On a 68040, this could be used in conjunction with an `asm` expression to access the `move16` instruction which requires 16-byte aligned operands.

You can also specify the alignment of structure fields. For example, to create a double-word aligned `int` pair, you could write:

```
struct foo { int x[2] __attribute__ ((aligned (8))); };
```

This is an alternative to creating a union with a `double` member that forces the union to be double-word aligned.

It is not possible to specify the alignment of functions; the alignment of functions is determined by the machine's requirements and cannot be changed. You cannot specify alignment for a typedef name because such a name is just an alias, not a distinct type.

As in the preceding examples, you can explicitly specify the alignment (in bytes) that you wish the compiler to use for a given variable or structure field. Alternatively, you can leave out the alignment factor and just ask the compiler to align a variable or field to the maximum useful alignment for the target machine you are compiling for. For example, you could write:

```
short array[3] __attribute__ ((aligned));
```

Whenever you leave out the alignment factor in an `aligned` attribute specification, the compiler automatically sets the alignment for the declared variable or field to the largest alignment which is ever used for any data type on the target machine you are compiling for. Doing this can often make copy operations more efficient, because the compiler can use whatever instructions copy the biggest chunks of memory when performing copies to or from the variables or fields that you have aligned this way.

The `aligned` attribute can only increase the alignment; but you can decrease it by specifying `packed` as well. See below.

Note that the effectiveness of `aligned` attributes may be limited by inherent limitations in your linker. On many systems, the linker is only able to arrange for variables to be aligned up to a certain maximum alignment. (For some linkers, the maximum supported alignment may be very very small.) If your linker is only able to align variables up to a maximum of 8 byte alignment, then specifying `aligned(16)` in an `__attribute__` will still only provide you with 8 byte alignment. See your linker documentation for further information.

`common` This attribute specifically requests GNU CC to make an uninitialised variable "common", instead of allocating space for it. This is only relevant if you have specified the '`-fno-common`', '`-mips16`' or '`-mips16e`' flag.

`mode (`*mode*`)`

This attribute specifies the data type for the declaration—whichever type corresponds to the mode *mode*. This in effect lets you request an integer or floating point type according to its width.

You may also specify a mode of '`byte`' or '`__byte__`' to indicate the mode corresponding to a one-byte integer, '`word`' or '`__word__`' for the mode of a one-word integer, and '`pointer`' or '`__pointer__`' for the mode used to represent pointers. The mode '`arg`' or '`__arg__`' represents the size of a stack argument slot or register.

`nocommon` This attribute specifically requests GNU CC not to make an uninitialised variable "common" but instead to allocate space for it directly. If you specify the '`-fno-common`' flag, GNU CC will do this for all variables.

Specifying the `nocommon` attribute for a variable provides an initialization of zeros. A variable may only be initialized in one source file.

`packed` The `packed` attribute specifies that a variable or structure field should have the smallest possible alignment—one byte for a variable, and one bit for a field, unless you specify a larger value with the `aligned` attribute.

Here is a structure in which the field `x` is packed, so that it immediately follows `a`:

```
struct foo
{
  char a;
  int x[2] __attribute__ ((packed));
};
```

`section ("section-name")`

Normally, the compiler places the objects it generates in sections like `data` and `bss`. Sometimes, however, you need additional sections, or you need certain particular variables to appear in special sections, for example to map to special hardware. The `section` attribute specifies that a variable (or function) lives in a particular section. For example, this small program uses several specific section names:

```
struct duart a __attribute__ ((section ("DUART_A"))) = { 0 };
struct duart b __attribute__ ((section ("DUART_B"))) = { 0 };
char stack[10000] __attribute__ ((section ("STACK"))) = { 0 };
int init_data __attribute__ ((section ("INITDATA"))) = 0;

main()
{
  /* Initialize stack pointer */
  init_sp (stack + sizeof (stack));

  /* Initialize initialized data */
  memcpy (&init_data, &data, &edata - &data);

  /* Turn on the serial ports */
  init_duart (&a);
  init_duart (&b);
}
```

Use the `section` attribute with an *initialized* definition of a *global* variable, as shown in the example. GNU CC issues a warning and otherwise ignores the `section` attribute in uninitialized variable declarations.

You may only use the `section` attribute with a fully initialized global definition because of the way linkers work. The linker requires each object be defined once, with the exception that uninitialized variables tentatively go in the `common` (or `bss`) section and can be multiply `"defined"`. You can force a variable to be initialized with the '`-fno-common`' flag or the `nocommon` attribute.

Some file formats do not support arbitrary sections so the `section` attribute is not available on all platforms. If you need to map the entire contents of a module to a particular section, consider using the facilities of the linker instead.

`transparent_union`

This attribute, attached to a function parameter which is a union, means that the corresponding argument may have the type of any union member, but the argument is passed as if its type were that of the first union member. For more details see See Section 3.30 [Type Attributes], page 95. You can also use this attribute on a `typedef` for a union data type; then it applies to all function parameters with that type.

`unused` This attribute, attached to a variable, means that the variable is meant to be possibly unused. GNU CC will not produce a warning for this variable.

`weak` The `weak` attribute is described in See Section 3.23 [Function Attributes], page 86.

To specify multiple attributes, separate them by commas within the double parentheses: for example, '`__attribute__ ((aligned (16), packed))`'.

3.30 Specifying Attributes of Types

The keyword `__attribute__` allows you to specify special attributes of `struct` and `union` types when you define such types. This keyword is followed by an attribute specification inside double parentheses. Three attributes are currently defined for types: `aligned`, `packed`, and `transparent_union`. Other attributes are defined for functions (see Section 3.23 [Function Attributes], page 86) and for variables (see Section 3.29 [Variable Attributes], page 92).

You may also specify any one of these attributes with '__' preceding and following its keyword. This allows you to use these attributes in header files without being concerned about a possible macro of the same name. For example, you may use `__aligned__` instead of `aligned`.

You may specify the `aligned` and `transparent_union` attributes either in a `typedef` declaration or just past the closing curly brace of a complete enum, struct or union type *definition* and the `packed` attribute only past the closing brace of a definition.

You may also specify attributes between the enum, struct or union tag and the name of the type rather than after the closing brace.

`aligned (`*alignment*`)`

This attribute specifies a minimum alignment (in bytes) for variables of the specified type. For example, the declarations:

```
struct S { short f[3]; } __attribute__ ((aligned (8)));
typedef int more_aligned_int __attribute__ ((aligned (8)));
```

force the compiler to insure (as far as it can) that each variable whose type is `struct S` or `more_aligned_int` will be allocated and aligned *at least* on a 8-byte boundary. On a Sparc, having all variables of type `struct S` aligned to 8-byte boundaries allows the compiler to use the `ldd` and `std` (doubleword load and store) instructions when copying one variable of type `struct S` to another, thus improving run-time efficiency.

Note that the alignment of any given `struct` or `union` type is required by the ANSI C standard to be at least a perfect multiple of the lowest common multiple of the alignments of all of the members of the `struct` or `union` in question. This means that you *can* effectively adjust the alignment of a `struct` or `union` type by attaching an `aligned` attribute to any one of the members of such a type, but the notation illustrated in the example above is a more obvious, intuitive, and readable way to request the compiler to adjust the alignment of an entire `struct` or `union` type.

As in the preceding example, you can explicitly specify the alignment (in bytes) that you wish the compiler to use for a given `struct` or `union` type. Alternatively, you can leave out the alignment factor and just ask the compiler to align a type to the maximum useful alignment for the target machine you are compiling for. For example, you could write:

```
struct S { short f[3]; } __attribute__ ((aligned));
```

Whenever you leave out the alignment factor in an `aligned` attribute specification, the compiler automatically sets the alignment for the type to the largest alignment which is ever used for any data type on the target machine you are compiling for. Doing this can often make copy operations more efficient, because the compiler can use whatever instructions copy the biggest chunks of memory when performing copies to or from the variables which have types that you have aligned this way.

In the example above, if the size of each `short` is 2 bytes, then the size of the entire `struct S` type is 6 bytes. The smallest power of two which is greater than or equal to that is 8, so the compiler sets the alignment for the entire `struct S` type to 8 bytes.

Note that although you can ask the compiler to select a time-efficient alignment for a given type and then declare only individual stand-alone objects of that type, the compiler's ability to select a time-efficient alignment is primarily useful only when you plan to create arrays of variables having the relevant (efficiently aligned) type. If you declare or use arrays of variables of an efficiently-aligned type, then it is likely that your program will also be doing pointer arithmetic (or subscripting, which amounts to the same thing) on pointers to the relevant type, and the code that the compiler generates for these pointer arithmetic operations will often be more efficient for efficiently-aligned types than for other types.

The `aligned` attribute can only increase the alignment; but you can decrease it by specifying `packed` as well. See below.

Note that the effectiveness of `aligned` attributes may be limited by inherent limitations in your linker. On many systems, the linker is only able to arrange for variables to be aligned up to a certain maximum alignment. (For some linkers, the maximum supported alignment may be very very small.) If your linker is only able to align variables up to a maximum of 8 byte alignment, then specifying `aligned(16)` in an `__attribute__` will still only provide you with 8 byte alignment. See your linker documentation for further information.

`packed` This attribute, attached to an `enum`, `struct`, or `union` type definition, specified that the minimum required memory be used to represent the type.

Specifying this attribute for `struct` and `union` types is equivalent to specifying the `packed` attribute on each of the structure or union members. Specifying the '`-fshort-enums`' flag on the line is equivalent to specifying the `packed` attribute on all `enum` definitions.

You may only specify this attribute after a closing curly brace on an `enum` definition, not in a `typedef` declaration, unless that declaration also contains the definition of the `enum`.

`transparent_union`

This attribute, attached to a `union` type definition, indicates that any function parameter having that union type causes calls to that function to be treated in a special way.

First, the argument corresponding to a transparent union type can be of any type in the union; no cast is required. Also, if the union contains a pointer type, the corresponding argument can be a null pointer constant or a void pointer expression; and if the union contains a void pointer type, the corresponding argument can be any pointer expres-

sion. If the union member type is a pointer, qualifiers like `const` on the referenced type must be respected, just as with normal pointer conversions.

Second, the argument is passed to the function using the calling conventions of first member of the transparent union, not the calling conventions of the union itself. All members of the union must have the same machine representation; this is necessary for this argument passing to work properly.

Transparent unions are designed for library functions that have multiple interfaces for compatibility reasons. For example, suppose the `wait` function must accept either a value of type `int *` to comply with Posix, or a value of type `union wait *` to comply with the 4.1BSD interface. If `wait`'s parameter were `void *`, `wait` would accept both kinds of arguments, but it would also accept any other pointer type and this would make argument type checking less useful. Instead, `<sys/wait.h>` might define the interface as follows:

```
typedef union
  {
    int *__ip;
    union wait *__up;
  } wait_status_ptr_t __attribute__ ((__transparent_union__));

pid_t wait (wait_status_ptr_t);
```

This interface allows either `int *` or `union wait *` arguments to be passed, using the `int *` calling convention. The program can call `wait` with arguments of either type:

```
int w1 () { int w; return wait (&w); }
int w2 () { union wait w; return wait (&w); }
```

With this interface, `wait`'s implementation might look like this:

```
pid_t wait (wait_status_ptr_t p)
{
  return waitpid (-1, p.__ip, 0);
}
```

`unused` When attached to a type (including a `union` or a `struct`), this attribute means that variables of that type are meant to appear possibly unused. GNU CC will not produce a warning for any variables of that type, even if the variable appears to do nothing. This is often the case with lock or thread classes, which are usually defined and then not referenced, but contain constructors and destructors that have nontrivial bookkeeping functions.

To specify multiple attributes, separate them by commas within the double parentheses: for example, '`__attribute__ ((aligned (16), packed))`'.

3.31 An Inline Function is As Fast As a Macro

By declaring a function `inline`, you can direct GNU CC to integrate that function's code into the code for its callers. This makes execution faster by eliminating the function-call overhead; in addition, if any of the actual argument values are constant, their known values may permit simplifications at compile time so that not all of the inline function's code needs to be included. The effect on code size is less predictable; object code may be larger or smaller with function inlining,

depending on the particular case. Inlining of functions is an optimization and it really "works" only in optimizing compilation. If you don't use '`-O`', no function is really inline.

To declare a function inline, use the `inline` keyword in its declaration, like this:

```
inline int
inc (int *a)
{
  (*a)++;
}
```

(If you are writing a header file to be included in ANSI C programs, write `__inline__` instead of `inline`. See Section 3.36 [Alternate Keywords], page 109.) You can also make all "simple enough" functions inline with the option '`-finline-functions`'.

Note that certain usages in a function definition can make it unsuitable for inline substitution. Among these usages are: use of varargs, use of alloca, use of variable sized data types (see Section 3.14 [Variable Length], page 81), use of computed goto (see Section 3.3 [Labels as Values], page 74), use of nonlocal goto, and nested functions (see Section 3.4 [Nested Functions], page 75). Using '`-Winline`' will warn when a function marked `inline` could not be substituted, and will give the reason for the failure.

Note that in C and Objective C, unlike C++, the `inline` keyword does not affect the linkage of the function.

GNU CC automatically inlines member functions defined within the class body of C++ programs even if they are not explicitly declared `inline`. (You can override this with '`-fno-default-inline`'; see Section 2.5 [Options Controlling C++ Dialect], page 15.)

When a function is both inline and `static`, if all calls to the function are integrated into the caller, and the function's address is never used, then the function's own assembler code is never referenced. In this case, GNU CC does not actually output assembler code for the function, unless you specify the option '`-fkeep-inline-functions`'. Some calls cannot be integrated for various reasons (in particular, calls that precede the function's definition cannot be integrated, and neither can recursive calls within the definition). If there is a nonintegrated call, then the function is compiled to assembler code as usual. The function must also be compiled as usual if the program refers to its address, because that can't be inlined.

When an inline function is not `static`, then the compiler must assume that there may be calls from other source files; since a global symbol can be defined only once in any program, the function must not be defined in the other source files, so the calls therein cannot be integrated. Therefore, a non-`static` inline function is always compiled on its own in the usual fashion.

If you specify both `inline` and `extern` in the function definition, then the definition is used only for inlining. In no case is the function compiled on its own, not even if you refer to its address explicitly. Such an address becomes an external reference, as if you had only declared the function, and had not defined it.

This combination of `inline` and `extern` has almost the effect of a macro. The way to use it is to put a function definition in a header file with these keywords, and put another copy of the definition (lacking `inline` and `extern`) in a library file. The definition in the header file will cause most calls to the function to be inlined. If any uses of the function remain, they will refer to the single copy in the library.

GNU C does not inline any functions when not optimizing. It is not clear whether it is better to inline or not, in this case, but we found that a correct implementation when not optimizing was difficult. So we did the easy thing, and turned it off.

3.32 Assembler Instructions with C Expression Operands

In an assembler instruction using `asm`, you can specify the operands of the instruction using C expressions. This means you need not guess which registers or memory locations will contain the data you want to use.

You must specify an assembler instruction template much like what appears in a machine description, plus an operand constraint string for each operand.

For example, here is how to use the 68881's `fsinx` instruction:

```
asm ("fsinx %1,%0" : "=f" (result) : "f" (angle));
```

Here `angle` is the C expression for the input operand while `result` is that of the output operand. Each has '`"f"`' as its operand constraint, saying that a floating point register is required. The '`=`' in '`=f`' indicates that the operand is an output; all output operands' constraints must use '`=`'. The constraints use the same language used in the machine description (see Section 3.33 [Constraints], page 102).

Each operand is described by an operand-constraint string followed by the C expression in parentheses. A colon separates the assembler template from the first output operand and another separates the last output operand from the first input, if any. Commas separate the operands within each group. The total number of operands is limited to ten or to the maximum number of operands in any instruction pattern in the machine description, whichever is greater.

If there are no output operands but there are input operands, you must place two consecutive colons surrounding the place where the output operands would go.

Output operand expressions must be lvalues; the compiler can check this. The input operands need not be lvalues. The compiler cannot check whether the operands have data types that are reasonable for the instruction being executed. It does not parse the assembler instruction template and does not know what it means or even whether it is valid assembler input. The extended `asm` feature is most often used for machine instructions the compiler itself does not know exist. If the output expression cannot be directly addressed (for example, it is a bit field), your constraint must allow a register. In that case, GNU CC will use the register as the output of the `asm`, and then store that register into the output.

The ordinary output operands must be write-only; GNU CC will assume that the values in these operands before the instruction are dead and need not be generated. Extended asm supports input-output or read-write operands. Use the constraint character '`+`' to indicate such an operand and list it with the output operands.

When the constraints for the read-write operand (or the operand in which only some of the bits are to be changed) allows a register, you may, as an alternative, logically split its function into two separate operands, one input operand and one write-only output operand. The connection between them is expressed by constraints which say they need to be in the same location when the instruction executes. You can use the same C expression for both operands, or different expressions. For example, here we write the (fictitious) '`combine`' instruction with `bar` as its read-only source operand and `foo` as its read-write destination:

```
asm ("combine %2,%0" : "=r" (foo) : "0" (foo), "g" (bar));
```

The constraint ‘"0"’ for operand 1 says that it must occupy the same location as operand 0. A digit in constraint is allowed only in an input operand and it must refer to an output operand.

Only a digit in the constraint can guarantee that one operand will be in the same place as another. The mere fact that `foo` is the value of both operands is not enough to guarantee that they will be in the same place in the generated assembler code. The following would not work reliably:

```
asm ("combine %2,%0" : "=r" (foo) : "r" (foo), "g" (bar));
```

Various optimizations or reloading could cause operands 0 and 1 to be in different registers; GNU CC knows no reason not to do so. For example, the compiler might find a copy of the value of `foo` in one register and use it for operand 1, but generate the output operand 0 in a different register (copying it afterward to `foo`’s own address). Of course, since the register for operand 1 is not even mentioned in the assembler code, the result will not work, but GNU CC can’t tell that.

Some instructions clobber specific hard registers. To describe this, write a third colon after the input operands, followed by the names of the clobbered hard registers (given as strings). Here is a realistic example for the VAX:

```
asm volatile ("movc3 %0,%1,%2"
              : /* no outputs */
              : "g" (from), "g" (to), "g" (count)
              : "r0", "r1", "r2", "r3", "r4", "r5");
```

You may not write a clobber description in a way that overlaps with an input or output operand. For example, you may not have an operand describing a register class with one member if you mention that register in the clobber list. There is no way for you to specify that an input operand is modified without also specifying it as an output operand. Note that if all the output operands you specify are for this purpose (and hence unused), you will then also need to specify `volatile` for the `asm` construct, as described below, to prevent GNU CC from deleting the `asm` statement as unused.

If you refer to a particular hardware register from the assembler code, you will probably have to list the register after the third colon to tell the compiler the register’s value is modified. In some assemblers, the register names begin with ‘`%`’; to produce one ‘`%`’ in the assembler code, you must write ‘`%%`’ in the input.

If your assembler instruction can alter the condition code register, add ‘`cc`’ to the list of clobbered registers. GNU CC on some machines represents the condition codes as a specific hardware register; ‘`cc`’ serves to name this register. On other machines, the condition code is handled differently, and specifying ‘`cc`’ has no effect. But it is valid no matter what the machine.

If your assembler instruction modifies memory in an unpredictable fashion, add ‘`memory`’ to the list of clobbered registers. This will cause GNU CC to not keep memory values cached in registers across the assembler instruction.

You can put multiple assembler instructions together in a single `asm` template, separated either with newlines (written as ‘`\n`’) or with semicolons if the assembler allows such semicolons. The GNU assembler allows semicolons and most Unix assemblers seem to do so. The input operands are guaranteed not to use any of the clobbered registers, and neither will the output operands’ addresses, so you can read and write the clobbered registers as many times as you like. Here is an example of multiple instructions in a template; it assumes the subroutine `_foo` accepts arguments in registers 9 and 10:

```
asm ("movl %0,r9;movl %1,r10;call _foo"
```

```
      : /* no outputs */
      : "g" (from), "g" (to)
      : "r9", "r10");
```

Unless an output operand has the ‘&’ constraint modifier, GNU CC may allocate it in the same register as an unrelated input operand, on the assumption the inputs are consumed before the outputs are produced. This assumption may be false if the assembler code actually consists of more than one instruction. In such a case, use ‘&’ for each output operand that may not overlap an input. See Section 3.33.3 [Modifiers], page 105.

If you want to test the condition code produced by an assembler instruction, you must include a branch and a label in the `asm` construct, as follows:

```
asm ("clr %0;frob %1;beq 0f;mov #1,%0;0:"
     : "g" (result)
     : "g" (input));
```

This assumes your assembler supports local labels, as the GNU assembler and most Unix assemblers do.

Speaking of labels, jumps from one `asm` to another are not supported. The compiler’s optimizers do not know about these jumps, and therefore they cannot take account of them when deciding how to optimize.

Usually the most convenient way to use these `asm` instructions is to encapsulate them in macros that look like functions. For example,

```
#define sin(x)       \
({ double __value, __arg = (x);   \
   asm ("fsinx %1,%0": "=f" (__value): "f" (__arg));  \
   __value; })
```

Here the variable `__arg` is used to make sure that the instruction operates on a proper `double` value, and to accept only those arguments `x` which can convert automatically to a `double`.

Another way to make sure the instruction operates on the correct data type is to use a cast in the `asm`. This is different from using a variable `__arg` in that it converts more different types. For example, if the desired type were `int`, casting the argument to `int` would accept a pointer with no complaint, while assigning the argument to an `int` variable named `__arg` would warn about using a pointer unless the caller explicitly casts it.

If an `asm` has output operands, GNU CC assumes for optimization purposes the instruction has no side effects except to change the output operands. This does not mean instructions with a side effect cannot be used, but you must be careful, because the compiler may eliminate them if the output operands aren’t used, or move them out of loops, or replace two with one if they constitute a common subexpression. Also, if your instruction does have a side effect on a variable that otherwise appears not to change, the old value of the variable may be reused later if it happens to be found in a register.

You can prevent an `asm` instruction from being deleted, moved significantly, or combined, by writing the keyword `volatile` after the `asm`. For example:

```
#define get_and_set_priority(new)  \
({ int __old; \
   asm volatile ("get_and_set_priority %0, %1": "=g" (__old) : "g" (new)); \
   __old; })
```

If you write an `asm` instruction with no outputs, GNU CC will know the instruction has side-effects and will not delete the instruction or move it outside of loops. If the side-effects of your instruction are not purely external, but will affect variables in your program in ways other than reading the inputs and clobbering the specified registers or memory, you should write the `volatile` keyword to prevent future versions of GNU CC from moving the instruction around within a core region.

An `asm` instruction without any operands or clobbers (and "old style" `asm`) will not be deleted or moved significantly, regardless, unless it is unreachable, the same wasy as if you had written a `volatile` keyword.

Note that even a volatile `asm` instruction can be moved in ways that appear insignificant to the compiler, such as across jump instructions. You can't expect a sequence of volatile `asm` instructions to remain perfectly consecutive. If you want consecutive output, use a single `asm`.

It is a natural idea to look for a way to give access to the condition code left by the assembler instruction. However, when we attempted to implement this, we found no way to make it work reliably. The problem is that output operands might need reloading, which would result in additional following "store" instructions. On most machines, these instructions would alter the condition code before there was time to test it. This problem doesn't arise for ordinary "test" and "compare" instructions because they don't have any output operands.

For reasons similar to those described above, it is not possible to give an assembler instruction access to the condition code left by previous instructions.

If you are writing a header file that should be includable in ANSI C programs, write `__asm__` instead of `asm`. See Section 3.36 [Alternate Keywords], page 109.

3.33 Constraints for `asm` Operands

Here are specific details on what constraint letters you can use with `asm` operands. 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.

3.33.1 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.

'`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.

'>' A memory operand with autoincrement addressing (either preincrement or postincrement) is allowed.

'r' A register operand is allowed provided that it is in a general register.

'd', 'a', 'f', ...
Other letters can be defined in machine-dependent fashion to stand for particular classes of registers. 'd', 'a' and 'f' are defined on the 68000/68020 to stand for data, address and floating point registers.

'i' An immediate integer operand (one with constant value) is allowed. This includes symbolic constants whose values will be known only at assembly time.

'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`) 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.

‘0’, ‘1’, ‘2’, ... ‘9’
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 is called a *matching constraint* and what it really means is that the assembler has only a single operand that fills two roles which `asm` distinguishes. For example, an add instruction uses two input operands and an output operand, 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.

‘p’ 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.

‘Q’, ‘R’, ‘S’, ... ‘U’
Letters in the range ‘Q’ through ‘U’ may be defined in a machine-dependent fashion to stand for arbitrary operand types.

3.33.2 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.

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.

3.33.3 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 interchange the two operands if that is the cheapest way to make all operands fit the constraints.

'`#`' 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.

3.33.4 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 3.33.1

[Simple Constraints], page 102), and 'I', usually the letter indicating the most common immediate-constant format.

For each machine architecture, the 'config/*machine*.h' file 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 interesting for asm. The constraints are defined through these macros:

REG_CLASS_FROM_LETTER
: Register class constraints (usually lower case).

CONST_OK_FOR_LETTER_P
: Immediate constant constraints, for non-floating point constants of word size or smaller precision (usually upper case).

CONST_DOUBLE_OK_FOR_LETTER_P
: Immediate constant constraints, for all floating point constants and for constants of greater than word size precision (usually upper case).

EXTRA_CONSTRAINT
: Special cases of registers or memory. This macro is not required, and is only defined for some machines.

Inspecting these macro definitions in the compiler source for your machine is the best way to be certain you have the right constraints. However, here is a summary of the machine-dependent constraints for MIPS CPUs.

MIPS—'mips.h'

d	General-purpose integer register
f	Floating-point register (if available)
h	'Hi' register
l	'Lo' register
x	'Hi' or 'Lo' register
y	General-purpose integer register
z	Floating-point status register
I	Signed 16 bit constant (for arithmetic instructions)
J	Zero
K	Zero-extended 16-bit constant (for logic instructions)
L	Constant with low 16 bits zero (can be loaded with lui)
M	32 bit constant which requires two instructions to load (a constant which is not 'I', 'K', or 'L')
N	Negative 16 bit constant
O	Exact power of two
P	Positive 16 bit constant

`G`	Floating point zero
`Q`	Memory reference that can be loaded with more than one instruction (‘`m`’ is preferable for `asm` statements)
`R`	Memory reference that can be loaded with one instruction (‘`m`’ is preferable for `asm` statements)
`S`	Memory reference in external OSF/rose PIC format (‘`m`’ is preferable for `asm` statements)

3.34 Controlling Names Used in Assembler Code

You can specify the name to be used in the assembler code for a C function or variable by writing the `asm` (or `__asm__`) keyword after the declarator as follows:

```
int foo asm ("myfoo") = 2;
```

This specifies that the name to be used for the variable `foo` in the assembler code should be ‘`myfoo`’ rather than the usual ‘`_foo`’.

On systems where an underscore is normally prepended to the name of a C function or variable, this feature allows you to define names for the linker that do not start with an underscore.

You cannot use `asm` in this way in a function *definition*; but you can get the same effect by writing a declaration for the function before its definition and putting `asm` there, like this:

```
extern func () asm ("FUNC");

func (x, y)
     int x, y;
...
```

It is up to you to make sure that the assembler names you choose do not conflict with any other assembler symbols. Also, you must not use a register name; that would produce completely invalid assembler code. GNU CC does not as yet have the ability to store static variables in registers. Perhaps that will be added.

3.35 Variables in Specified Registers

GNU C allows you to put a few global variables into specified hardware registers. You can also specify the register in which an ordinary register variable should be allocated.

- Global register variables reserve registers throughout the program. This may be useful in programs such as programming language interpreters which have a couple of global variables that are accessed very often.
- Local register variables in specific registers do not reserve the registers. The compiler’s data flow analysis is capable of determining where the specified registers contain live values, and where they are available for other uses. Stores into local register variables may be deleted when they appear to be dead according to dataflow analysis. References to local register variables may be deleted or moved or simplified.

  These local variables are sometimes convenient for use with the extended `asm` feature (see Section 3.32 [Extended Asm], page 99), if you want to write one output of the assembler instruction directly into a particular register. (This will work provided the register you specify fits the constraints specified for that operand in the `asm`.)

3.35.1 Defining Global Register Variables

You can define a global register variable in GNU C like this:

```
register int *foo asm ("a3");
```

Here `a3` is the name of the register which should be used. Choose a register which is normally saved and restored by function calls on your machine, so that library routines will not clobber it.

Naturally the register name is cpu-dependent, so you would need to conditionalize your program according to cpu type. The register `a3` would be a good choice on a 68000 for a variable of pointer type. On machines with register windows, be sure to choose a "global" register that is not affected magically by the function call mechanism.

In addition, operating systems on one type of cpu may differ in how they name the registers; then you would need additional conditionals. For example, some 68000 operating systems call this register `%a3`.

Eventually there may be a way of asking the compiler to choose a register automatically, but first we need to figure out how it should choose and how to enable you to guide the choice. No solution is evident.

Defining a global register variable in a certain register reserves that register entirely for this use, at least within the current compilation. The register will not be allocated for any other purpose in the functions in the current compilation. The register will not be saved and restored by these functions. Stores into this register are never deleted even if they would appear to be dead, but references may be deleted or moved or simplified.

It is not safe to access the global register variables from signal handlers, or from more than one thread of control, because the system library routines may temporarily use the register for other things (unless you recompile them specially for the task at hand).

It is not safe for one function that uses a global register variable to call another such function `foo` by way of a third function `lose` that was compiled without knowledge of this variable (i.e. in a different source file in which the variable wasn't declared). This is because `lose` might save the register and put some other value there. For example, you can't expect a global register variable to be available in the comparison-function that you pass to `qsort`, since `qsort` might have put something else in that register. (If you are prepared to recompile `qsort` with the same global register variable, you can solve this problem.)

If you want to recompile `qsort` or other source files which do not actually use your global register variable, so that they will not use that register for any other purpose, then it suffices to specify the compiler option '`-ffixed-`*reg*'. You need not actually add a global register declaration to their source code.

A function which can alter the value of a global register variable cannot safely be called from a function compiled without this variable, because it could clobber the value the caller expects to find there on return. Therefore, the function which is the entry point into the part of the program that uses the global register variable must explicitly save and restore the value which belongs to its caller.

On most machines, `longjmp` will restore to each global register variable the value it had at the time of the `setjmp`. On some machines, however, `longjmp` will not change the value of global register variables. To be portable, the function that called `setjmp` should make other arrangements to save the values of the global register variables, and to restore them in a `longjmp`. This way, the same thing will happen regardless of what `longjmp` does.

All global register variable declarations must precede all function definitions. If such a declaration could appear after function definitions, the declaration would be too late to prevent the register from being used for other purposes in the preceding functions.

Global register variables may not have initial values, because an executable file has no means to supply initial contents for a register.

On MIPS, s0 . . . s7 should be suitable. Of course, it will not do to use more than a few of those. For MIPS16 it would also not be wise to use s0 or s1, since that would severely restrict the number of registers available to the compiler and generate very poor code.

3.35.2 Specifying Registers for Local Variables

You can define a local register variable with a specified register like this:

```
register int *foo asm ("a3");
```

Here `a3` is the name of the register which should be used. Note that this is the same syntax used for defining global register variables, but for a local variable it would appear within a function.

Naturally the register name is cpu-dependent, but this is not a problem, since specific registers are most often useful with explicit assembler instructions (see Section 3.32 [Extended Asm], page 99). Both of these things generally require that you conditionalize your program according to cpu type.

In addition, operating systems on one type of cpu may differ in how they name the registers; then you would need additional conditionals. For example, some 68000 operating systems call this register `%a3`.

Defining such a register variable does not reserve the register; it remains available for other uses in places where flow control determines the variable's value is not live. However, these registers are made unavailable for use in the reload pass; excessive use of this feature leaves the compiler too few available registers to compile certain functions.

This option does not guarantee that GNU CC will generate code that has this variable in the register you specify at all times. You may not code an explicit reference to this register in an `asm` statement and assume it will always refer to this variable.

Stores into local register variables may be deleted when they appear to be dead according to dataflow analysis. References to local register variables may be deleted or moved or simplified.

3.36 Alternate Keywords

The option '`-traditional`' disables certain keywords; '`-ansi`' disables certain others. This causes trouble when you want to use GNU C extensions, or ANSI C features, in a general-purpose header file that should be usable by all programs, including ANSI C programs and traditional ones. The keywords `asm`, `typeof` and `inline` cannot be used since they won't work in a program compiled with '`-ansi`', while the keywords `const`, `volatile`, `signed`, `typeof` and `inline` won't work in a program compiled with '`-traditional`'.

The way to solve these problems is to put '__' at the beginning and end of each problematical keyword. For example, use `__asm__` instead of `asm`, `__const__` instead of `const`, and `__inline__` instead of `inline`.

Other C compilers won't accept these alternative keywords; if you want to compile with another compiler, you can define the alternate keywords as macros to replace them with the customary keywords. It looks like this:

```
#ifndef __GNUC__
#define __asm__ asm
#endif
```

‘`-pedantic`’ and other options cause warnings for many GNU C extensions. You can prevent such warnings within one expression by writing `__extension__` before the expression. `__extension__` has no effect aside from this.

3.37 Incomplete `enum` Types

You can define an `enum` tag without specifying its possible values. This results in an incomplete type, much like what you get if you write `struct foo` without describing the elements. A later declaration which does specify the possible values completes the type.

You can’t allocate variables or storage using the type while it is incomplete. However, you can work with pointers to that type.

This extension may not be very useful, but it makes the handling of `enum` more consistent with the way `struct` and `union` are handled.

This extension is not supported by GNU C++.

3.38 Function Names as Strings

GNU CC predefines two magic identifiers to hold the name of the current function. The identifier `__FUNCTION__` holds the name of the function as it appears in the source. The identifier `__PRETTY_FUNCTION__` holds the name of the function pretty printed in a language specific fashion.

These names are always the same in a C function, but in a C++ function they may be different. For example, this program:

```
extern "C" {
extern int printf (char *, ...);
}

class a {
 public:
  sub (int i)
    {
      printf ("__FUNCTION__ = %s\n", __FUNCTION__);
      printf ("__PRETTY_FUNCTION__ = %s\n", __PRETTY_FUNCTION__);
    }
};

int
main (void)
{
  a ax;
  ax.sub (0);
  return 0;
}
```

gives this output:

```
__FUNCTION__ = sub
__PRETTY_FUNCTION__ = int  a::sub (int)
```

The compiler automagically replaces the identifiers with a string literal containing the appropriate name. Thus, they are neither preprocessor macros, like `__FILE__` and `__LINE__`, nor variables. This means that they catenate with other string literals, and that they can be used to initialize char arrays. For example

```
char here[] = "Function " __FUNCTION__ " in " __FILE__;
```

On the other hand, '`#ifdef __FUNCTION__`' does not have any special meaning inside a function, since the preprocessor does not do anything special with the identifier `__FUNCTION__`.

GNU CC also supports the magic word `__func__`, defined by the ISO standard C-99:

> The identifier `__func__` is implicitly declared by the translator as if, immediately following the opening brace of each function definition, the declaration
>
> ```
> static const char __func__[] = "function-name";
> ```
>
> appeared, where function-name is the name of the lexically-enclosing function. This name is the unadorned name of the function.

By this definition, `__func__` is a variable, not a string literal. In particular, `__func__` does not catenate with other string literals.

In `C++`, `__FUNCTION__` and `__PRETTY_FUNCTION__` are variables, declared in the same way as `__func__`.

3.39 Getting the Return or Frame Address of a Function

These functions may be used to get information about the callers of a function.

`__builtin_return_address` (*level*)

> This function returns the return address of the current function, or of one of its callers. The *level* argument is number of frames to scan up the call stack. A value of `0` yields the return address of the current function, a value of `1` yields the return address of the caller of the current function, and so forth.
>
> The *level* argument must be a constant integer.
>
> On some machines it may be impossible to determine the return address of any function other than the current one; in such cases, or when the top of the stack has been reached, this function will return `0`.
>
> This function should only be used with a non-zero argument for debugging purposes.

`__builtin_frame_address` (*level*)

> This function is similar to `__builtin_return_address`, but it returns the address of the function frame rather than the return address of the function. Calling `__builtin_frame_address` with a value of `0` yields the frame address of the current function, a value of `1` yields the frame address of the caller of the current function, and so forth.
>
> The frame is the area on the stack which holds local variables and saved registers. The frame address is normally the address of the first word pushed on to the stack by the

function. However, the exact definition depends upon the processor and the calling convention. If the processor has a dedicated frame pointer register, and the function has a frame, then `__builtin_frame_address` will return the value of the frame pointer register.

The caveats that apply to `__builtin_return_address` apply to this function as well.

3.40 Other built-in functions provided by GNU CC

GNU CC provides a large number of built-in functions other than the ones mentioned above. Some of these are for internal use in the processing of exceptions or variable-length argument lists and will not be documented here because they may change from time to time; we do not recommend general use of these functions.

The remaining functions are provided for optimization purposes.

GNU CC includes builtin versions of many of the functions in the standard C library. These will always be treated as having the same meaning as the C library function even if you specify the '`-fno-builtin`' (see Section 2.4 [C Dialect Options], page 11) option. These functions correspond to the C library functions `abort`, `abs`, `alloca`, `cos`, `cosf`, `cosl`, `exit`, `_exit`, `fabs`, `fabsf`, `fabsl`, `ffs`, `labs`, `memcmp`, `memcpy`, `memset`, `sin`, `sinf`, `sinl`, `sqrt`, `sqrtf`, `sqrtl`, `strcmp`, `strcpy`, and `strlen`.

`__builtin_constant_p (`*exp*`)`

You can use the builtin function `__builtin_constant_p` to determine if a value is known to be constant at compile-time and hence that GNU CC can perform constant-folding on expressions involving that value. The argument of the function is the value to test. The function returns the integer 1 if the argument is known to be a compile-time constant and 0 if it is not known to be a compile-time constant. A return of 0 does not indicate that the value is *not* a constant, but merely that GNU CC cannot prove it is a constant with the specified value of the '`-O`' option.

You would typically use this function in an embedded application where memory was a critical resource. If you have some complex calculation, you may want it to be folded if it involves constants, but need to call a function if it does not. For example:

```
#define Scale_Value(X)  \
  (__builtin_constant_p (X) ? ((X) * SCALE + OFFSET) : Scale (X))
```

You may use this builtin function in either a macro or an inline function. However, if you use it in an inlined function and pass an argument of the function as the argument to the builtin, GNU CC will never return 1 when you call the inline function with a string constant or constructor expression (see Section 3.19 [Constructors], page 84) and will not return 1 when you pass a constant numeric value to the inline function unless you specify the '`-O`' option.

`__builtin_expect(`*exp*`, `*c*`)`

You may use `__builtin_expect` to provide the compiler with branch prediction information. In general, you should prefer to use actual profile feedback for this ('`-fprofile-arcs`'), as programmers are notoriously bad at predicting how their programs actually perform. However, there are applications in which this data is hard to collect.

The return value is the value of *exp*, which should be an integral expression. The value of *c* must be a compile-time constant. The semantics of the builtin are that it is expected that *exp* == *c*. For example:

```
if (__builtin_expect (x, 0))
  foo ();
```

would indicate that we do not expect to call `foo`, since we expect `x` to be zero. Since you are limited to integral expressions for *exp*, you should use constructions such as

```
if (__builtin_expect (ptr != NULL, 1))
  error ();
```

when testing pointer or floating-point values.

3.41 Redefining typedef names

GNU CC allows you to redefine a typedef name, as long as the previous type and the new type are the same. ISO/ANSI does not allow typedefs to be redefined.

3.42 Deprecated Features

In the past, the GNU C++ compiler was extended to experiment with new features, at a time when the C++ language was still evolving. Now that the C++ standard is complete, some of those features are superseded by superior alternatives. Using the old features might cause a warning in some cases that the feature will be dropped in the future. In other cases, the feature might be gone already.

While the list below is not exhaustive, it documents some of the options that are now deprecated:

`-fexternal-templates`
`-falt-external-templates`

These are two of the many ways for g++ to implement template instantiation. See Section 4.6 [Template Instantiation], page 120. The C++ standard clearly defines how template definitions have to be organized across implementation units. g++ has an implicit instantiation mechanism that should work just fine for standard-conforming code.

4 Extensions to the C++ Language

The GNU compiler provides these extensions to the C++ language (and you can also use most of the C language extensions in your C++ programs). If you want to write code that checks whether these features are available, you can test for the GNU compiler the same way as for C programs: check for a predefined macro `__GNUC__`. You can also use `__GNUG__` to test specifically for GNU C++ (see section "Standard Predefined Macros" in *The C Preprocessor*).

4.1 Named Return Values in C++

GNU C++ extends the function-definition syntax to allow you to specify a name for the result of a function outside the body of the definition, in C++ programs:

```
type
functionname (args) return resultname;
{
  ...
  body
  ...
}
```

You can use this feature to avoid an extra constructor call when a function result has a class type. For example, consider a function `m`, declared as '`X v = m ();`', whose result is of class `X`:

```
X
m ()
{
  X b;
  b.a = 23;
  return b;
}
```

Although `m` appears to have no arguments, in fact it has one implicit argument: the address of the return value. At invocation, the address of enough space to hold `v` is sent in as the implicit argument. Then `b` is constructed and its `a` field is set to the value 23. Finally, a copy constructor (a constructor of the form '`X(X&)`') is applied to `b`, with the (implicit) return value location as the target, so that `v` is now bound to the return value.

But this is wasteful. The local `b` is declared just to hold something that will be copied right out. While a compiler that combined an "elision" algorithm with interprocedural data flow analysis could conceivably eliminate all of this, it is much more practical to allow you to assist the compiler in generating efficient code by manipulating the return value explicitly, thus avoiding the local variable and copy constructor altogether.

Using the extended GNU C++ function-definition syntax, you can avoid the temporary allocation and copying by naming `r` as your return value at the outset, and assigning to its `a` field directly:

```
X
m () return r;
{
  r.a = 23;
}
```

The declaration of `r` is a standard, proper declaration, whose effects are executed **before** any of the body of `m`.

Functions of this type impose no additional restrictions; in particular, you can execute `return` statements, or return implicitly by reaching the end of the function body ("falling off the edge"). Cases like

```
X
m () return r (23);
{
  return;
}
```

(or even '`X m () return r (23); { }`') are unambiguous, since the return value `r` has been initialized in either case. The following code may be hard to read, but also works predictably:

```
X
m () return r;
{
  X b;
  return b;
}
```

The return value slot denoted by `r` is initialized at the outset, but the statement '`return b;`' overrides this value. The compiler deals with this by destroying `r` (calling the destructor if there is one, or doing nothing if there is not), and then reinitializing `r` with `b`.

This extension is provided primarily to help people who use overloaded operators, where there is a great need to control not just the arguments, but the return values of functions. For classes where the copy constructor incurs a heavy performance penalty (especially in the common case where there is a quick default constructor), this is a major savings. The disadvantage of this extension is that you do not control when the default constructor for the return value is called: it is always called at the beginning.

4.2 Minimum and Maximum Operators in C++

It is very convenient to have operators which return the "minimum" or the "maximum" of two arguments. In GNU C++ (but not in GNU C),

a `<?` *b* is the *minimum*, returning the smaller of the numeric values *a* and *b*;

a `>?` *b* is the *maximum*, returning the larger of the numeric values *a* and *b*.

These operations are not primitive in ordinary C++, since you can use a macro to return the minimum of two things in C++, as in the following example.

```
#define MIN(X,Y) ((X) < (Y) ? : (X) : (Y))
```

You might then use '`int min = MIN (i, j);`' to set *min* to the minimum value of variables *i* and *j*.

However, side effects in `X` or `Y` may cause unintended behavior. For example, `MIN (i++, j++)` will fail, incrementing the smaller counter twice. A GNU C extension allows you to write safe macros that avoid this kind of problem (see Section 3.6 [Naming an Expression's Type], page 78). However, writing `MIN` and `MAX` as macros also forces you to use function-call notation for a fundamental arithmetic operation. Using GNU C++ extensions, you can write '`int min = i <? j;`' instead.

Since `<?` and `>?` are built into the compiler, they properly handle expressions with side-effects; '`int min = i++ <? j++;`' works correctly.

4.3 When is a Volatile Object Accessed?

Both the C and C++ standard have the concept of volatile objects. These are normally accessed by pointers and used for accessing hardware. The standards encourage compilers to refrain from optimizations on concerning accesses to volatile objects that it might perform on non-volatile objects. The C standard leaves it implementation defined as to what constitutes a volatile access. The C++ standard omits to specify this, except to say that C++ should behave in a similar manner to C with respect to volatiles, where possible. The minimum either standard specifies is that at a sequence point all previous access to volatile objects have stabilized and no subsequent accesses have occurred. Thus an implementation is free to reorder and combine volatile accesses which occur between sequence points, but cannot do so for accesses across a sequence point. The use of volatiles does not allow you to violate the restriction on updating objects multiple times within a sequence point.

In most expressions, it is intuitively obvious what is a read and what is a write. For instance

```
volatile int *dst = <somevalue>;
volatile int *src = <someothervalue>;
*dst = *src;
```

will cause a read of the volatile object pointed to by *src* and stores the value into the volatile object pointed to by *dst*. There is no guarantee that these reads and writes are atomic, especially for objects larger than `int`.

Less obvious expressions are where something which looks like an access is used in a void context. An example would be,

```
volatile int *src = <somevalue>;
*src;
```

With C, such expressions are rvalues, and as rvalues cause a read of the object, gcc interprets this as a read of the volatile being pointed to. The C++ standard specifies that such expressions do not undergo lvalue to rvalue conversion, and that the type of the dereferenced object may be incomplete. The C++ standard does not specify explicitly that it is this lvalue to rvalue conversion which is responsible for causing an access. However, there is reason to believe that it is, because otherwise certain simple expressions become undefined. However, because it would surprise most programmers, g++ treats dereferencing a pointer to volatile object of complete type in a void context as a read of the object. When the object has incomplete type, g++ issues a warning.

```
struct S;
struct T {int m;};
volatile S *ptr1 = <somevalue>;
volatile T *ptr2 = <somevalue>;
*ptr1;
*ptr2;
```

In this example, a warning is issued for `*ptr1`, and `*ptr2` causes a read of the object pointed to. If you wish to force an error on the first case, you must force a conversion to rvalue with, for instance a static cast, `static_cast<S>(*ptr1)`.

When using a reference to volatile, g++ does not treat equivalent expressions as accesses to volatiles, but instead issues a warning that no volatile is accessed. The rationale for this is that otherwise it becomes difficult to determine where volatile access occur, and not possible to ignore the return value from functions returning volatile references. Again, if you wish to force a read, cast the reference to an rvalue.

4.4 Restricting Pointer Aliasing

As with gcc, g++ understands the C9X proposal of restricted pointers, specified with the __restrict__, or __restrict type qualifier. Because you cannot compile C++ by specifying the -flang-isoc9x language flag, restrict is not a keyword in C++.

In addition to allowing restricted pointers, you can specify restricted references, which indicate that the reference is not aliased in the local context.

```
void fn (int *__restrict__ rptr, int &__restrict__ rref)
{
  ...
}
```

In the body of fn, *rptr* points to an unaliased integer and *rref* refers to a (different) unaliased integer.

You may also specify whether a member function's *this* pointer is unaliased by using __restrict__ as a member function qualifier.

```
void T::fn () __restrict__
{
  ...
}
```

Within the body of T::fn, *this* will have the effective definition T *__restrict__ const this. Notice that the interpretation of a __restrict__ member function qualifier is different to that of const or volatile qualifier, in that it is applied to the pointer rather than the object. This is consistent with other compilers which implement restricted pointers.

As with all outermost parameter qualifiers, __restrict__ is ignored in function definition matching. This means you only need to specify __restrict__ in a function definition, rather than in a function prototype as well.

4.5 Declarations and Definitions in One Header

C++ object definitions can be quite complex. In principle, your source code will need two kinds of things for each object that you use across more than one source file. First, you need an *interface* specification, describing its structure with type declarations and function prototypes. Second, you need the *implementation* itself. It can be tedious to maintain a separate interface description in a header file, in parallel to the actual implementation. It is also dangerous, since separate interface and implementation definitions may not remain parallel.

With GNU C++, you can use a single header file for both purposes.

> *Warning:* The mechanism to specify this is in transition. For the nonce, you must use one of two #pragma commands; in a future release of GNU C++, an alternative mechanism will make these #pragma commands unnecessary.

The header file contains the full definitions, but is marked with '#pragma interface' in the source code. This allows the compiler to use the header file only as an interface specification when ordinary source files incorporate it with #include. In the single source file where the full implementation belongs, you can use either a naming convention or '#pragma implementation' to indicate this alternate use of the header file.

`#pragma interface`

`#pragma interface "`*subdir/objects*`.h"`

Use this directive in *header files* that define object classes, to save space in most of the object files that use those classes. Normally, local copies of certain information (backup copies of inline member functions, debugging information, and the internal tables that implement virtual functions) must be kept in each object file that includes class definitions. You can use this pragma to avoid such duplication. When a header file containing '`#pragma interface`' is included in a compilation, this auxiliary information will not be generated (unless the main input source file itself uses '`#pragma implementation`'). Instead, the object files will contain references to be resolved at link time.

The second form of this directive is useful for the case where you have multiple headers with the same name in different directories. If you use this form, you must specify the same string to '`#pragma implementation`'.

`#pragma implementation`

`#pragma implementation "`*objects*`.h"`

Use this pragma in a *main input file*, when you want full output from included header files to be generated (and made globally visible). The included header file, in turn, should use '`#pragma interface`'. Backup copies of inline member functions, debugging information, and the internal tables used to implement virtual functions are all generated in implementation files.

If you use '`#pragma implementation`' with no argument, it applies to an include file with the same basename[1] as your source file. For example, in '`allclass.cc`', giving just '`#pragma implementation`' by itself is equivalent to '`#pragma implementation "allclass.h"`'.

In versions of GNU C++ prior to 2.6.0 '`allclass.h`' was treated as an implementation file whenever you would include it from '`allclass.cc`' even if you never specified '`#pragma implementation`'. This was deemed to be more trouble than it was worth, however, and disabled.

If you use an explicit '`#pragma implementation`', it must appear in your source file *before* you include the affected header files.

Use the string argument if you want a single implementation file to include code from multiple header files. (You must also use '`#include`' to include the header file; '`#pragma implementation`' only specifies how to use the file—it doesn't actually include it.)

There is no way to split up the contents of a single header file into multiple implementation files.

'`#pragma implementation`' and '`#pragma interface`' also have an effect on function inlining.

If you define a class in a header file marked with '`#pragma interface`', the effect on a function defined in that class is similar to an explicit `extern` declaration—the compiler emits no code at all to define an independent version of the function. Its definition is used only for inlining with its callers.

[1] A file's *basename* was the name stripped of all leading path information and of trailing suffixes, such as '`.h`' or '`.C`' or '`.cc`'.

Conversely, when you include the same header file in a main source file that declares it as '`#pragma implementation`', the compiler emits code for the function itself; this defines a version of the function that can be found via pointers (or by callers compiled without inlining). If all calls to the function can be inlined, you can avoid emitting the function by compiling with '`-fno-implement-inlines`'. If any calls were not inlined, you will get linker errors.

4.6 Where's the Template?

C++ templates are the first language feature to require more intelligence from the environment than one usually finds on a UNIX system. Somehow the compiler and linker have to make sure that each template instance occurs exactly once in the executable if it is needed, and not at all otherwise. There are two basic approaches to this problem, which I will refer to as the Borland model and the Cfront model.

Borland model
: Borland C++ solved the template instantiation problem by adding the code equivalent of common blocks to their linker; the compiler emits template instances in each translation unit that uses them, and the linker collapses them together. The advantage of this model is that the linker only has to consider the object files themselves; there is no external complexity to worry about. This disadvantage is that compilation time is increased because the template code is being compiled repeatedly. Code written for this model tends to include definitions of all templates in the header file, since they must be seen to be instantiated.

Cfront model
: The AT&T C++ translator, Cfront, solved the template instantiation problem by creating the notion of a template repository, an automatically maintained place where template instances are stored. A more modern version of the repository works as follows: As individual object files are built, the compiler places any template definitions and instantiations encountered in the repository. At link time, the link wrapper adds in the objects in the repository and compiles any needed instances that were not previously emitted. The advantages of this model are more optimal compilation speed and the ability to use the system linker; to implement the Borland model a compiler vendor also needs to replace the linker. The disadvantages are vastly increased complexity, and thus potential for error; for some code this can be just as transparent, but in practice it can been very difficult to build multiple programs in one directory and one program in multiple directories. Code written for this model tends to separate definitions of non-inline member templates into a separate file, which should be compiled separately.

When used with GNU ld version 2.8 or later on an ELF system such as Linux/GNU or Solaris 2, or on Microsoft Windows, g++ supports the Borland model. On other systems, g++ implements neither automatic model.

A future version of g++ will support a hybrid model whereby the compiler will emit any instantiations for which the template definition is included in the compile, and store template definitions and instantiation context information into the object file for the rest. The link wrapper will extract that information as necessary and invoke the compiler to produce the remaining instantiations. The linker will then combine duplicate instantiations.

In the mean time, you have the following options for dealing with template instantiations:

1. Compile your template-using code with '`-frepo`'. The compiler will generate files with the extension '`.rpo`' listing all of the template instantiations used in the corresponding object files which could be instantiated there; the link wrapper, '`collect2`', will then update the '`.rpo`' files to tell the compiler where to place those instantiations and rebuild any affected object files. The link-time overhead is negligible after the first pass, as the compiler will continue to place the instantiations in the same files.

   This is your best option for application code written for the Borland model, as it will just work. Code written for the Cfront model will need to be modified so that the template definitions are available at one or more points of instantiation; usually this is as simple as adding `#include <tmethods.cc>` to the end of each template header.

   For library code, if you want the library to provide all of the template instantiations it needs, just try to link all of its object files together; the link will fail, but cause the instantiations to be generated as a side effect. Be warned, however, that this may cause conflicts if multiple libraries try to provide the same instantiations. For greater control, use explicit instantiation as described in the next option.

2. Compile your code with '`-fno-implicit-templates`' to disable the implicit generation of template instances, and explicitly instantiate all the ones you use. This approach requires more knowledge of exactly which instances you need than do the others, but it's less mysterious and allows greater control. You can scatter the explicit instantiations throughout your program, perhaps putting them in the translation units where the instances are used or the translation units that define the templates themselves; you can put all of the explicit instantiations you need into one big file; or you can create small files like

   ```
   #include "Foo.h"
   #include "Foo.cc"

   template class Foo<int>;
   template ostream& operator <<
                   (ostream&, const Foo<int>&);
   ```

   for each of the instances you need, and create a template instantiation library from those.

   If you are using Cfront-model code, you can probably get away with not using '`-fno-implicit-templates`' when compiling files that don't '`#include`' the member template definitions.

   If you use one big file to do the instantiations, you may want to compile it without '`-fno-implicit-templates`' so you get all of the instances required by your explicit instantiations (but not by any other files) without having to specify them as well.

   g++ has extended the template instantiation syntax outlined in the Working Paper to allow forward declaration of explicit instantiations and instantiation of the compiler support data for a template class (i.e. the vtable) without instantiating any of its members:

   ```
   extern template int max (int, int);
   inline template class Foo<int>;
   ```

3. Do nothing. Pretend g++ does implement automatic instantiation management. Code written for the Borland model will work fine, but each translation unit will contain instances of each of the templates it uses. In a large program, this can lead to an unacceptable amount of code duplication.

4. Add '`#pragma interface`' to all files containing template definitions. For each of these files, add '`#pragma implementation "`*filename*`"`' to the top of some '`.C`' file which '`#include`'s it. Then compile everything with '`-fexternal-templates`'. The templates will then only be expanded in the translation unit which implements them (i.e. has a '`#pragma implementation`' line for the file where they live); all other files will use external references. If you're lucky, everything should work properly. If you get undefined symbol errors, you need to make sure that each template instance which is used in the program is used in the file which implements that template. If you don't have any use for a particular instance in that file, you can just instantiate it explicitly, using the syntax from the latest C++ working paper:

   ```
   template class A<int>;
   template ostream& operator << (ostream&, const A<int>&);
   ```

   This strategy will work with code written for either model. If you are using code written for the Cfront model, the file containing a class template and the file containing its member templates should be implemented in the same translation unit.

   A slight variation on this approach is to instead use the flag '`-falt-external-templates`'; this flag causes template instances to be emitted in the translation unit that implements the header where they are first instantiated, rather than the one which implements the file where the templates are defined. This header must be the same in all translation units, or things are likely to break.

   See Section 4.5 [Declarations and Definitions in One Header], page 118, for more discussion of these pragmas.

4.7 Extracting the function pointer from a bound pointer to member function

In C++, pointer to member functions (PMFs) are implemented using a wide pointer of sorts to handle all the possible call mechanisms; the PMF needs to store information about how to adjust the '`this`' pointer, and if the function pointed to is virtual, where to find the vtable, and where in the vtable to look for the member function. If you are using PMFs in an inner loop, you should really reconsider that decision. If that is not an option, you can extract the pointer to the function that would be called for a given object/PMF pair and call it directly inside the inner loop, to save a bit of time.

Note that you will still be paying the penalty for the call through a function pointer; on most modern architectures, such a call defeats the branch prediction features of the CPU. This is also true of normal virtual function calls.

The syntax for this extension is

```
extern A a;
extern int (A::*fp)();
typedef int (*fptr)(A *);

fptr p = (fptr)(a.*fp);
```

For PMF constants (i.e. expressions of the form '`&Klasse::Member`'), no object is needed to obtain the address of the function. They can be converted to function pointers directly:

```
fptr p1 = (fptr)(&A::foo);
```

You must specify '`-Wno-pmf-conversions`' to use this extension.

5 gcov: a Test Coverage Program

gcov is a tool you can use in conjunction with GNU CC to test code coverage in your programs.

This chapter describes version 1.5 of gcov.

5.1 Introduction to gcov

gcov is a test coverage program. Use it in concert with GNU CC to analyze your programs to help create more efficient, faster running code. You can use gcov as a profiling tool to help discover where your optimization efforts will best affect your code. You can also use gcov along with the other profiling tool, gprof, to assess which parts of your code use the greatest amount of computing time.

Profiling tools help you analyze your code's performance. Using a profiler such as gcov or gprof, you can find out some basic performance statistics, such as:

- how often each line of code executes
- what lines of code are actually executed
- how much computing time each section of code uses

Once you know these things about how your code works when compiled, you can look at each module to see which modules should be optimized. gcov helps you determine where to work on optimization.

Software developers also use coverage testing in concert with testsuites, to make sure software is actually good enough for a release. Testsuites can verify that a program works as expected; a coverage program tests to see how much of the program is exercised by the testsuite. Developers can then determine what kinds of test cases need to be added to the testsuites to create both better testing and a better final product.

You should compile your code without optimization if you plan to use gcov because the optimization, by combining some lines of code into one function, may not give you as much information as you need to look for 'hot spots' where the code is using a great deal of computer time. Likewise, because gcov accumulates statistics by line (at the lowest resolution), it works best with a programming style that places only one statement on each line. If you use complicated macros that expand to loops or to other control structures, the statistics are less helpful—they only report on the line where the macro call appears. If your complex macros behave like functions, you can replace them with inline functions to solve this problem.

gcov creates a logfile called '*sourcefile*.gcov' which indicates how many times each line of a source file '*sourcefile*.c' has executed. You can use these logfiles along with gprof to aid in fine-tuning the performance of your programs. gprof gives timing information you can use along with the information you get from gcov.

gcov works only on code compiled with GNU CC. It is not compatible with any other profiling or test coverage mechanism.

5.2 Invoking gcov

```
gcov [-b] [-c] [-v] [-n] [-l] [-f] [-o directory] sourcefile
```

`-b` Write branch frequencies to the output file, and write branch summary info to the standard output. This option allows you to see how often each branch in your program was taken.

`-c` Write branch frequencies as the number of branches taken, rather than the percentage of branches taken.

`-v` Display the `gcov` version number (on the standard error stream).

`-n` Do not create the `gcov` output file.

`-l` Create long file names for included source files. For example, if the header file '`x.h`' contains code, and was included in the file '`a.c`', then running `gcov` on the file '`a.c`' will produce an output file called '`a.c.x.h.gcov`' instead of '`x.h.gcov`'. This can be useful if '`x.h`' is included in multiple source files.

`-f` Output summaries for each function in addition to the file level summary.

`-o` The directory where the object files live. Gcov will search for `.bb`, `.bbg`, and `.da` files in this directory.

When using `gcov`, you must first compile your program with two special GNU CC options: '`-fprofile-arcs -ftest-coverage`'. This tells the compiler to generate additional information needed by gcov (basically a flow graph of the program) and also includes additional code in the object files for generating the extra profiling information needed by gcov. These additional files are placed in the directory where the source code is located.

Running the program will cause profile output to be generated. For each source file compiled with -fprofile-arcs, an accompanying `.da` file will be placed in the source directory.

Running `gcov` with your program's source file names as arguments will now produce a listing of the code along with frequency of execution for each line. For example, if your program is called '`tmp.c`', this is what you see when you use the basic `gcov` facility:

```
$ gcc -fprofile-arcs -ftest-coverage tmp.c
$ a.out
$ gcov tmp.c
 87.50% of 8 source lines executed in file tmp.c
Creating tmp.c.gcov.
```

The file '`tmp.c.gcov`' contains output from `gcov`. Here is a sample:

```
                main()
                {
             1    int i, total;

             1    total = 0;

            11    for (i = 0; i < 10; i++)
            10      total += i;

             1    if (total != 45)
        ######      printf ("Failure\n");
                  else
             1      printf ("Success\n");
             1  }
```

When you use the ‘-b’ option, your output looks like this:

```
$ gcov -b tmp.c
 87.50% of 8 source lines executed in file tmp.c
 80.00% of 5 branches executed in file tmp.c
 80.00% of 5 branches taken at least once in file tmp.c
 50.00% of 2 calls executed in file tmp.c
Creating tmp.c.gcov.
```

Here is a sample of a resulting ‘tmp.c.gcov’ file:

```
                main()
                {
            1     int i, total;

            1     total = 0;

           11     for (i = 0; i < 10; i++)
branch 0 taken = 91%
branch 1 taken = 100%
branch 2 taken = 100%
           10       total += i;

            1     if (total != 45)
branch 0 taken = 100%
      ######       printf ("Failure\n");
call 0 never executed
branch 1 never executed
                  else
            1       printf ("Success\n");
call 0 returns = 100%
            1   }
```

For each basic block, a line is printed after the last line of the basic block describing the branch or call that ends the basic block. There can be multiple branches and calls listed for a single source line if there are multiple basic blocks that end on that line. In this case, the branches and calls are each given a number. There is no simple way to map these branches and calls back to source constructs. In general, though, the lowest numbered branch or call will correspond to the leftmost construct on the source line.

For a branch, if it was executed at least once, then a percentage indicating the number of times the branch was taken divided by the number of times the branch was executed will be printed. Otherwise, the message “never executed” is printed.

For a call, if it was executed at least once, then a percentage indicating the number of times the call returned divided by the number of times the call was executed will be printed. This will usually be 100%, but may be less for functions call `exit` or `longjmp`, and thus may not return everytime they are called.

The execution counts are cumulative. If the example program were executed again without removing the `.da` file, the count for the number of times each line in the source was executed would be added to the results of the previous run(s). This is potentially useful in several ways. For example,

it could be used to accumulate data over a number of program runs as part of a test verification suite, or to provide more accurate long-term information over a large number of program runs.

The data in the `.da` files is saved immediately before the program exits. For each source file compiled with -fprofile-arcs, the profiling code first attempts to read in an existing `.da` file; if the file doesn't match the executable (differing number of basic block counts) it will ignore the contents of the file. It then adds in the new execution counts and finally writes the data to the file.

5.3 Using `gcov` with GCC Optimization

If you plan to use `gcov` to help optimize your code, you must first compile your program with two special GNU CC options: '`-fprofile-arcs -ftest-coverage`'. Aside from that, you can use any other GNU CC options; but if you want to prove that every single line in your program was executed, you should not compile with optimization at the same time. On some machines the optimizer can eliminate some simple code lines by combining them with other lines. For example, code like this:

```
if (a != b)
  c = 1;
else
  c = 0;
```

can be compiled into one instruction on some machines. In this case, there is no way for `gcov` to calculate separate execution counts for each line because there isn't separate code for each line. Hence the `gcov` output looks like this if you compiled the program with optimization:

```
      100  if (a != b)
      100    c = 1;
      100  else
      100    c = 0;
```

The output shows that this block of code, combined by optimization, executed 100 times. In one sense this result is correct, because there was only one instruction representing all four of these lines. However, the output does not indicate how many times the result was 0 and how many times the result was 1.

5.4 Brief description of `gcov` data files

`gcov` uses three files for doing profiling. The names of these files are derived from the original *source* file by substituting the file suffix with either `.bb`, `.bbg`, or `.da`. All of these files are placed in the same directory as the source file, and contain data stored in a platform-independent method.

The `.bb` and `.bbg` files are generated when the source file is compiled with the GNU CC '`-ftest-coverage`' option. The `.bb` file contains a list of source files (including headers), functions within those files, and line numbers corresponding to each basic block in the source file.

The `.bb` file format consists of several lists of 4-byte integers which correspond to the line numbers of each basic block in the file. Each list is terminated by a line number of 0. A line number of -1 is used to designate that the source file name (padded to a 4-byte boundary and followed by another -1) follows. In addition, a line number of -2 is used to designate that the name of a function (also padded to a 4-byte boundary and followed by a -2) follows.

The `.bbg` file is used to reconstruct the program flow graph for the source file. It contains a list of the program flow arcs (possible branches taken from one basic block to another) for each function which, in combination with the `.bb` file, enables gcov to reconstruct the program flow.

In the .bbg file, the format is:

```
number of basic blocks for function #0 (4-byte number)
total number of arcs for function #0 (4-byte number)
count of arcs in basic block #0 (4-byte number)
destination basic block of arc #0 (4-byte number)
flag bits (4-byte number)
destination basic block of arc #1 (4-byte number)
flag bits (4-byte number)
...
destination basic block of arc #N (4-byte number)
flag bits (4-byte number)
count of arcs in basic block #1 (4-byte number)
destination basic block of arc #0 (4-byte number)
flag bits (4-byte number)
...
```

A -1 (stored as a 4-byte number) is used to separate each function's list of basic blocks, and to verify that the file has been read correctly.

The .da file is generated when a program containing object files built with the GNU CC '-fprofile-arcs' option is executed. A separate .da file is created for each source file compiled with this option, and the name of the .da file is stored as an absolute pathname in the resulting object file. This path name is derived from the source file name by substituting a .da suffix.

The format of the .da file is fairly simple. The first 8-byte number is the number of counts in the file, followed by the counts (stored as 8-byte numbers). Each count corresponds to the number of times each arc in the program is executed. The counts are cumulative; each time the program is executed, it attemps to combine the existing .da files with the new counts for this invocation of the program. It ignores the contents of any .da files whose number of arcs doesn't correspond to the current program, and merely overwrites them instead.

All three of these files use the functions in gcov-io.h to store integers; the functions in this header provide a machine-independent mechanism for storing and retrieving data from a stream.

6 Known Causes of Trouble with GCC

This section describes known problems that affect users of GCC. Most of these are not GCC bugs per se—if they were, we would fix them. But the result for a user may be like the result of a bug.

Some of these problems are due to bugs in other software, some are missing features that are too much work to add, and some are places where people's opinions differ as to what is best.

6.1 Actual Bugs We Haven't Fixed Yet

- There are several obscure case of mis-using struct, union, and enum tags that are not detected as errors by the compiler.
- When '`-pedantic-errors`' is specified, GCC will incorrectly give an error message when a function name is specified in an expression involving the comma operator.
- Loop unrolling doesn't work properly for certain C++ programs. This is a bug in the C++ front end. It sometimes emits incorrect debug info, and the loop unrolling code is unable to recover from this error.

6.2 Problems Compiling Certain Programs

Certain programs have problems compiling.

- Parse errors may occur compiling X11 on a Decstation running Ultrix 4.2 because of problems in DEC's versions of the X11 header files '`X11/Xlib.h`' and '`X11/Xutil.h`'. People recommend adding '`-I/usr/include/mit`' to use the MIT versions of the header files, using the '`-traditional`' switch to turn off ANSI C, or fixing the header files by adding this:

  ```
  #ifdef __STDC__
  #define NeedFunctionPrototypes 0
  #endif
  ```

- On various 386 Unix systems derived from System V, including SCO, ISC, and ESIX, you may get error messages about running out of virtual memory while compiling certain programs.

  You can prevent this problem by linking GCC with the GNU malloc (which thus replaces the malloc that comes with the system). GNU malloc is available as a separate package, and also in the file '`src/gmalloc.c`' in the GNU Emacs 19 distribution.

  If you have installed GNU malloc as a separate library package, use this option when you relink GCC:

  ```
  MALLOC=/usr/local/lib/libgmalloc.a
  ```

  Alternatively, if you have compiled '`gmalloc.c`' from Emacs 19, copy the object file to '`gmalloc.o`' and use this option when you relink GCC:

  ```
  MALLOC=gmalloc.o
  ```

6.3 Incompatibilities of GCC

There are several noteworthy incompatibilities between GNU C and K&R (non-ANSI) versions of C. The '`-traditional`' option eliminates many of these incompatibilities, *but not all,* by telling GNU C to behave like a K&R C compiler.

- GCC normally makes string constants read-only. If several identical-looking string constants are used, GCC stores only one copy of the string.

  One consequence is that you cannot call `mktemp` with a string constant argument. The function `mktemp` always alters the string its argument points to.

  Another consequence is that `sscanf` does not work on some systems when passed a string constant as its format control string or input. This is because `sscanf` incorrectly tries to write into the string constant. Likewise `fscanf` and `scanf`.

  The best solution to these problems is to change the program to use `char`-array variables with initialization strings for these purposes instead of string constants. But if this is not possible, you can use the '`-fwritable-strings`' flag, which directs GCC to handle string constants the same way most C compilers do. '`-traditional`' also has this effect, among others.

- `-2147483648` is positive.

  This is because 2147483648 cannot fit in the type `int`, so (following the ANSI C rules) its data type is `unsigned long int`. Negating this value yields 2147483648 again.

- GCC does not substitute macro arguments when they appear inside of string constants. For example, the following macro in GCC

  ```
  #define foo(a) "a"
  ```

  will produce output `"a"` regardless of what the argument *a* is.

  The '`-traditional`' option directs GCC to handle such cases (among others) in the old-fashioned (non-ANSI) fashion.

- When you use `setjmp` and `longjmp`, the only automatic variables guaranteed to remain valid are those declared `volatile`. This is a consequence of automatic register allocation. Consider this function:

  ```
  jmp_buf j;

  foo ()
  {
    int a, b;

    a = fun1 ();
    if (setjmp (j))
      return a;

    a = fun2 ();
    /* longjmp (j) may occur in fun3. */
    return a + fun3 ();
  }
  ```

  Here `a` may or may not be restored to its first value when the `longjmp` occurs. If `a` is allocated in a register, then its first value is restored; otherwise, it keeps the last value stored in it.

  If you use the '`-W`' option with the '`-O`' option, you will get a warning when GCC thinks such a problem might be possible.

  The '`-traditional`' option directs GNU C to put variables in the stack by default, rather than in registers, in functions that call `setjmp`. This results in the behavior found in traditional C compilers.

- Programs that use preprocessing directives in the middle of macro arguments do not work with GCC. For example, a program like this will not work:

```
foobar (
#define luser
        hack)
```

  ANSI C does not permit such a construct. It would make sense to support it when '`-traditional`' is used, but it is too much work to implement.

- Declarations of external variables and functions within a block apply only to the block containing the declaration. In other words, they have the same scope as any other declaration in the same place.

  In some other C compilers, a `extern` declaration affects all the rest of the file even if it happens within a block.

  The '`-traditional`' option directs GNU C to treat all `extern` declarations as global, like traditional compilers.

- In traditional C, you can combine `long`, etc., with a typedef name, as shown here:

```
typedef int foo;
typedef long foo bar;
```

  In ANSI C, this is not allowed: `long` and other type modifiers require an explicit `int`. Because this criterion is expressed by Bison grammar rules rather than C code, the '`-traditional`' flag cannot alter it.

- PCC allows typedef names to be used as function parameters. The difficulty described immediately above applies here too.

- PCC allows whitespace in the middle of compound assignment operators such as '`+=`'. GCC, following the ANSI standard, does not allow this. The difficulty described immediately above applies here too.

- GCC complains about unterminated character constants inside of preprocessing conditionals that fail. Some programs have English comments enclosed in conditionals that are guaranteed to fail; if these comments contain apostrophes, GCC will probably report an error. For example, this code would produce an error:

```
#if 0
You can't expect this to work.
#endif
```

  The best solution to such a problem is to put the text into an actual C comment delimited by '`/*...*/`'. However, '`-traditional`' suppresses these error messages.

- Many user programs contain the declaration '`long time ();`'. In the past, the system header files on many systems did not actually declare `time`, so it did not matter what type your program declared it to return. But in systems with ANSI C headers, `time` is declared to return `time_t`, and if that is not the same as `long`, then '`long time ();`' is erroneous.

  The solution is to change your program to use `time_t` as the return type of `time`.

- When compiling functions that return `float`, PCC converts it to a double. GCC actually returns a `float`. If you are concerned with PCC compatibility, you should declare your functions to return `double`; you might as well say what you mean.

- When compiling functions that return structures or unions, GCC output code normally uses a method different from that used on most versions of Unix. As a result, code compiled with GCC cannot call a structure-returning function compiled with PCC, and vice versa.

  The method used by GCC is as follows: a structure or union which is 1, 2, 4 or 8 bytes long is returned like a scalar. A structure or union with any other size is stored into an address supplied by the caller (usually in a special, fixed register, but on some machines it is passed on the stack). The machine-description macros `STRUCT_VALUE` and `STRUCT_INCOMING_VALUE` tell 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. GCC does not use this method because it is slower and nonreentrant.

  On some newer machines, PCC uses a reentrant convention for all structure and union returning. GCC on most of these machines uses a compatible convention when returning structures and unions in memory, but still returns small structures and unions in registers.

  You can tell GCC to use a compatible convention for all structure and union returning with the option '`-fpcc-struct-return`'.

- GNU C complains about program fragments such as '`0x74ae-0x4000`' which appear to be two hexadecimal constants separated by the minus operator. Actually, this string is a single *preprocessing token*. Each such token must correspond to one token in C. Since this does not, GNU C prints an error message. Although it may appear obvious that what is meant is an operator and two values, the ANSI C standard specifically requires that this be treated as erroneous.

  A *preprocessing token* is a *preprocessing number* if it begins with a digit and is followed by letters, underscores, digits, periods and '`e+`', '`e-`', '`E+`', or '`E-`' character sequences.

  To make the above program fragment valid, place whitespace in front of the minus sign. This whitespace will end the preprocessing number.

6.4 Standard Libraries

GCC by itself attempts to be what the ISO/ANSI C standard calls a *conforming freestanding implementation*. This means all ANSI C language features are available. If you're building for MIPS targets using MIPS Technologies' MIPS SDE package, then you will also have a compatible and standards-conformant C library and matching include files.

6.5 Disappointments and Misunderstandings

These problems are perhaps regrettable, but we don't know any practical way around them.

- Certain local variables aren't recognized by debuggers when you compile with optimization.

  This occurs because sometimes GCC optimizes the variable out of existence. There is no way to tell the debugger how to compute the value such a variable "would have had", and it is not clear that would be desirable anyway. So GCC simply does not mention the eliminated variable when it writes debugging information.

  You have to expect a certain amount of disagreement between the executable and your source code, when you use optimization.

- Users often think it is a bug when GCC reports an error for code like this:

```
int foo (struct mumble *);

struct mumble { ... };

int foo (struct mumble *x)
{ ... }
```

  This code really is erroneous, because the scope of `struct mumble` in the prototype is limited to the argument list containing it. It does not refer to the `struct mumble` defined with file scope immediately below—they are two unrelated types with similar names in different scopes.

  But in the definition of `foo`, the file-scope type is used because that is available to be inherited. Thus, the definition and the prototype do not match, and you get an error.

  This behavior may seem silly, but it's what the ANSI standard specifies. It is easy enough for you to make your code work by moving the definition of `struct mumble` above the prototype. It's not worth being incompatible with ANSI C just to avoid an error for the example shown above.

- Accesses to bitfields even in volatile objects works by accessing larger objects, such as a byte or a word. You cannot rely on what size of object is accessed in order to read or write the bitfield; it may even vary for a given bitfield according to the precise usage.

  If you care about controlling the amount of memory that is accessed, use volatile but do not use bitfields.

- On the MIPS, variable argument functions using '`varargs.h`' cannot have a floating point value for the first argument. The reason for this is that in the absence of a prototype in scope, if the first argument is a floating point, it is passed in a floating point register, rather than an integer register.

  If the code is rewritten to use the ANSI standard '`stdarg.h`' method of variable arguments, and the prototype is in scope at the time of the call, everything will work fine.

6.6 Common Misunderstandings with GNU C++

C++ is a complex language and an evolving one, and its standard definition (the ISO C++ standard) was only recently completed. As a result, your C++ compiler may occasionally surprise you, even when its behavior is correct. This section discusses some areas that frequently give rise to questions of this sort.

6.6.1 Declare *and* Define Static Members

When a class has static data members, it is not enough to *declare* the static member; you must also *define* it. For example:

```
class Foo
{
  ...
  void method();
  static int bar;
};
```

This declaration only establishes that the class `Foo` has an `int` named `Foo::bar`, and a member function named `Foo::method`. But you still need to define *both* `method` and `bar` elsewhere. According to the draft ANSI standard, you must supply an initializer in one (and only one) source file, such as:

```
int Foo::bar = 0;
```

Other C++ compilers may not correctly implement the standard behavior. As a result, when you switch to `g++` from one of these compilers, you may discover that a program that appeared to work correctly in fact does not conform to the standard: `g++` reports as undefined symbols any static data members that lack definitions.

6.6.2 Temporaries May Vanish Before You Expect

It is dangerous to use pointers or references to *portions* of a temporary object. The compiler may very well delete the object before you expect it to, leaving a pointer to garbage. The most common place where this problem crops up is in classes like string classes, especially ones that define a conversion function to type `char *` or `const char *` – which is one reason why the standard `string` class requires you to call the `c_str` member function. However, any class that returns a pointer to some internal structure is potentially subject to this problem.

For example, a program may use a function `strfunc` that returns `string` objects, and another function `charfunc` that operates on pointers to `char`:

```
string strfunc ();
void charfunc (const char *);

void
f ()
{
  const char *p = strfunc().c_str();
  ...
  charfunc (p);
  ...
  charfunc (p);
}
```

In this situation, it may seem reasonable to save a pointer to the C string returned by the `c_str` member function and use that rather than call `c_str` repeatedly. However, the temporary string created by the call to `strfunc` is destroyed after `p` is initialized, at which point `p` is left pointing to freed memory.

Code like this may run successfully under some other compilers, particularly obsolete cfront-based compilers that delete temporaries along with normal local variables. However, the GNU C++ behavior is standard-conforming, so if your program depends on late destruction of temporaries it is not portable.

The safe way to write such code is to give the temporary a name, which forces it to remain until the end of the scope of the name. For example:

```
string& tmp = strfunc ();
charfunc (tmp.c_str ());
```

6.6.3 Implicit Copy-Assignment for Virtual Bases

When a base class is virtual, only one subobject of the base class belongs to each full object. Also, the constructors and destructors are invoked only once, and called from the most-derived class. However, such objects behave unspecified when being assigned. For example:

```
struct Base{
  char *name;
  Base(char *n) : name(strdup(n)){}
  Base& operator= (const Base& other){
   free (name);
   name = strdup (other.name);
  }
};

struct A:virtual Base{
  int val;
  A():Base("A"){}
};

struct B:virtual Base{
  int bval;
  B():Base("B"){}
};

struct Derived:public A, public B{
  Derived():Base("Derived"){}
};

void func(Derived &d1, Derived &d2)
{
  d1 = d2;
}
```

The C++ standard specifies that ‘`Base::Base`’ is only called once when constructing or copy-constructing a Derived object. It is unspecified whether ‘`Base::operator=`’ is called more than once when the implicit copy-assignment for Derived objects is invoked (as it is inside ‘`func`’ in the example).

g++ implements the `"intuitive"` algorithm for copy-assignment: assign all direct bases, then assign all members. In that algorithm, the virtual base subobject can be encountered many times. In the example, copying proceeds in the following order: ‘`val`’, ‘`name`’ (via `strdup`), ‘`bval`’, and ‘`name`’ again.

If application code relies on copy-assignment, a user-defined copy-assignment operator removes any uncertainties. With such an operator, the application can define whether and how the virtual base subobject is assigned.

6.7 Caveats of using `protoize`

The conversion programs `protoize` and `unprotoize` can sometimes change a source file in a way that won’t work unless you rearrange it.

- `protoize` can insert references to a type name or type tag before the definition, or in a file where they are not defined.

  If this happens, compiler error messages should show you where the new references are, so fixing the file by hand is straightforward.
- There are some C constructs which `protoize` cannot figure out. For example, it can't determine argument types for declaring a pointer-to-function variable; this you must do by hand. `protoize` inserts a comment containing '`???`' each time it finds such a variable; so you can find all such variables by searching for this string. ANSI C does not require declaring the argument types of pointer-to-function types.
- Using `unprotoize` can easily introduce bugs. If the program relied on prototypes to bring about conversion of arguments, these conversions will not take place in the program without prototypes. One case in which you can be sure `unprotoize` is safe is when you are removing prototypes that were made with `protoize`; if the program worked before without any prototypes, it will work again without them.

  You can find all the places where this problem might occur by compiling the program with the '`-Wconversion`' option. It prints a warning whenever an argument is converted.
- Both conversion programs can be confused if there are macro calls in and around the text to be converted. In other words, the standard syntax for a declaration or definition must not result from expanding a macro. This problem is inherent in the design of C and cannot be fixed. If only a few functions have confusing macro calls, you can easily convert them manually.
- `protoize` cannot get the argument types for a function whose definition was not actually compiled due to preprocessing conditionals. When this happens, `protoize` changes nothing in regard to such a function. `protoize` tries to detect such instances and warn about them.

  You can generally work around this problem by using `protoize` step by step, each time specifying a different set of '`-D`' options for compilation, until all of the functions have been converted. There is no automatic way to verify that you have got them all, however.
- Confusion may result if there is an occasion to convert a function declaration or definition in a region of source code where there is more than one formal parameter list present. Thus, attempts to convert code containing multiple (conditionally compiled) versions of a single function header (in the same vicinity) may not produce the desired (or expected) results.

  If you plan on converting source files which contain such code, it is recommended that you first make sure that each conditionally compiled region of source code which contains an alternative function header also contains at least one additional follower token (past the final right parenthesis of the function header). This should circumvent the problem.
- `unprotoize` can become confused when trying to convert a function definition or declaration which contains a declaration for a pointer-to-function formal argument which has the same name as the function being defined or declared. We recommand you avoid such choices of formal parameter names.
- You might also want to correct some of the indentation by hand and break long lines. (The conversion programs don't write lines longer than eighty characters in any case.)

6.8 Certain Changes We Don't Want to Make

This section lists changes that people frequently request, but which we do not make because we think GCC is better without them.

- Checking the number and type of arguments to a function which has an old-fashioned definition and no prototype.

  Such a feature would work only occasionally—only for calls that appear in the same file as the called function, following the definition. The only way to check all calls reliably is to add a prototype for the function. But adding a prototype eliminates the motivation for this feature. So the feature is not worthwhile.

- Warning about using an expression whose type is signed as a shift count.

  Shift count operands are probably signed more often than unsigned. Warning about this would cause far more annoyance than good.

- Warning about assigning a signed value to an unsigned variable.

  Such assignments must be very common; warning about them would cause more annoyance than good.

- Warning when a non-void function value is ignored.

  Coming as I do from a Lisp background, I balk at the idea that there is something dangerous about discarding a value. There are functions that return values which some callers may find useful; it makes no sense to clutter the program with a cast to `void` whenever the value isn't useful.

- Assuming (for optimization) that the address of an external symbol is never zero.

  This assumption is false on certain systems when '`#pragma weak`' is used.

- Making '`-fshort-enums`' the default.

  This would cause storage layout to be incompatible with most other C compilers. And it doesn't seem very important, given that you can get the same result in other ways. The case where it matters most is when the enumeration-valued object is inside a structure, and in that case you can specify a field width explicitly.

- Making bitfields unsigned by default on particular machines where "the ABI standard" says to do so.

  The ANSI C standard leaves it up to the implementation whether a bitfield declared plain `int` is signed or not. This in effect creates two alternative dialects of C.

  The GNU C compiler supports both dialects; you can specify the signed dialect with '`-fsigned-bitfields`' and the unsigned dialect with '`-funsigned-bitfields`'. However, this leaves open the question of which dialect to use by default.

  Currently, the preferred dialect makes plain bitfields signed, because this is simplest. Since `int` is the same as `signed int` in every other context, it is cleanest for them to be the same in bitfields as well.

  Some computer manufacturers have published Application Binary Interface standards which specify that plain bitfields should be unsigned. It is a mistake, however, to say anything about this issue in an ABI. This is because the handling of plain bitfields distinguishes two dialects of C. Both dialects are meaningful on every type of machine. Whether a particular object file was compiled using signed bitfields or unsigned is of no concern to other object files, even if they access the same bitfields in the same data structures.

  A given program is written in one or the other of these two dialects. The program stands a chance to work on most any machine if it is compiled with the proper dialect. It is unlikely to work at all if compiled with the wrong dialect.

Many users appreciate the GNU C compiler because it provides an environment that is uniform across machines. These users would be inconvenienced if the compiler treated plain bitfields differently on certain machines.

Occasionally users write programs intended only for a particular machine type. On these occasions, the users would benefit if the GNU C compiler were to support by default the same dialect as the other compilers on that machine. But such applications are rare. And users writing a program to run on more than one type of machine cannot possibly benefit from this kind of compatibility.

This is why GCC does and will treat plain bitfields in the same fashion on all types of machines (by default).

There are some arguments for making bitfields unsigned by default on all machines. If, for example, this becomes a universal de facto standard, it would make sense for GCC to go along with it. This is something to be considered in the future.

(Of course, users strongly concerned about portability should indicate explicitly in each bitfield whether it is signed or not. In this way, they write programs which have the same meaning in both C dialects.)

- Undefining `__STDC__` when '`-ansi`' is not used.

  Currently, GCC defines `__STDC__` as long as you don't use '`-traditional`'. This provides good results in practice.

  Programmers normally use conditionals on `__STDC__` to ask whether it is safe to use certain features of ANSI C, such as function prototypes or ANSI token concatenation. Since plain '`gcc`' supports all the features of ANSI C, the correct answer to these questions is "yes".

  Some users try to use `__STDC__` to check for the availability of certain library facilities. This is actually incorrect usage in an ANSI C program, because the ANSI C standard says that a conforming freestanding implementation should define `__STDC__` even though it does not have the library facilities. '`gcc -ansi -pedantic`' is a conforming freestanding implementation, and it is therefore required to define `__STDC__`, even though it does not come with an ANSI C library.

  Sometimes people say that defining `__STDC__` in a compiler that does not completely conform to the ANSI C standard somehow violates the standard. This is illogical. The standard is a standard for compilers that claim to support ANSI C, such as '`gcc -ansi`'—not for other compilers such as plain '`gcc`'. Whatever the ANSI C standard says is relevant to the design of plain '`gcc`' without '`-ansi`' only for pragmatic reasons, not as a requirement.

  GCC normally defines `__STDC__` to be 1, and in addition defines `__STRICT_ANSI__` if you specify the '`-ansi`' option. On some hosts, system include files use a different convention, where `__STDC__` is normally 0, but is 1 if the user specifies strict conformance to the C Standard. GCC follows the host convention when processing system include files, but when processing user files it follows the usual GNU C convention.

- Undefining `__STDC__` in C++.

  Programs written to compile with C++-to-C translators get the value of `__STDC__` that goes with the C compiler that is subsequently used. These programs must test `__STDC__` to determine what kind of C preprocessor that compiler uses: whether they should concatenate tokens in the ANSI C fashion or in the traditional fashion.

  These programs work properly with GNU C++ if `__STDC__` is defined. They would not work otherwise.

In addition, many header files are written to provide prototypes in ANSI C but not in traditional C. Many of these header files can work without change in C++ provided `__STDC__` is defined. If `__STDC__` is not defined, they will all fail, and will all need to be changed to test explicitly for C++ as well.

- Deleting "empty" loops.

  Historically, GCC has not deleted "empty" loops under the assumption that the most likely reason you would put one in a program is to have a delay, so deleting them will not make real programs run any faster.

  However, the rationale here is that optimization of a nonempty loop cannot produce an empty one, which holds for C but is not always the case for C++.

  Moreover, with '`-funroll-loops`' small "empty" loops are already removed, so the current behavior is both sub-optimal and inconsistent and will change in the future.

- Making side effects happen in the same order as in some other compiler.

  It is never safe to depend on the order of evaluation of side effects. For example, a function call like this may very well behave differently from one compiler to another:

  ```
  void func (int, int);

  int i = 2;
  func (i++, i++);
  ```

  There is no guarantee (in either the C or the C++ standard language definitions) that the increments will be evaluated in any particular order. Either increment might happen first. `func` might get the arguments '`2, 3`', or it might get '`3, 2`', or even '`2, 2`'.

- Not allowing structures with volatile fields in registers.

  Strictly speaking, there is no prohibition in the ANSI C standard against allowing structures with volatile fields in registers, but it does not seem to make any sense and is probably not what you wanted to do. So the compiler will give an error message in this case.

6.9 Warning Messages and Error Messages

The GNU compiler can produce two kinds of diagnostics: errors and warnings. Each kind has a different purpose:

Errors report problems that make it impossible to compile your program. GCC reports errors with the source file name and line number where the problem is apparent.

Warnings report other unusual conditions in your code that *may* indicate a problem, although compilation can (and does) proceed. Warning messages also report the source file name and line number, but include the text '`warning:`' to distinguish them from error messages.

Warnings may indicate danger points where you should check to make sure that your program really does what you intend; or the use of obsolete features; or the use of nonstandard features of GNU C or C++. Many warnings are issued only if you ask for them, with one of the '`-W`' options (for instance, '`-Wall`' requests a variety of useful warnings).

GCC always tries to compile your program if possible; it never gratuitously rejects a program whose meaning is clear merely because (for instance) it fails to conform to a standard. In some cases, however, the C and C++ standards specify that certain extensions are forbidden, and a diagnostic *must* be issued by a conforming compiler. The '`-pedantic`' option tells GCC to issue warnings

in such cases; ‘`-pedantic-errors`’ says to make them errors instead. This does not mean that *all* non-ANSI constructs get warnings or errors.

See Section 2.6 [Options to Request or Suppress Warnings], page 21, for more detail on these and related command-line options.

7 Reporting Bugs

Apart from this introduction, most of this chapter is from the generic GNU C compiler manual. The chapter's advice on how to recognise and report bugs is valuable, but MIPS SDE users with current support should send reports of bugs to MIPS Technologies at `support@mips.com`. If you don't have current support and don't want to renew it, you can send a just-for-information report to MIPS Technologies, or follow the advice in this chapter and post to the global community, or both.

Your bug reports play an essential role in making GCC reliable.

When you encounter a problem, the first thing to do is to see if it is already known. See Chapter 6 [Trouble], page 131. If it isn't known, then you should report the problem.

Reporting a bug may help you by bringing a solution to your problem, or it may not. (If it does not, look in the service directory; see Chapter 8 [Service], page 151.) In any case, the principal function of a bug report is to help the entire community by making the next version of GCC work better. Bug reports are your contribution to the maintenance of GCC.

Since the maintainers are very overloaded, we cannot respond to every bug report. However, if the bug has not been fixed, we are likely to send you a patch and ask you to tell us whether it works.

In order for a bug report to serve its purpose, you must include the information that makes for fixing the bug.

7.1 Have You Found a Bug?

If you are not sure whether you have found a bug, here are some guidelines:

- If the compiler gets a fatal signal, for any input whatever, that is a compiler bug. Reliable compilers never crash.
- If the compiler produces invalid assembly code, for any input whatever (except an `asm` statement), that is a compiler bug, unless the compiler reports errors (not just warnings) which would ordinarily prevent the assembler from being run.
- If the compiler produces valid assembly code that does not correctly execute the input source code, that is a compiler bug.

  However, you must double-check to make sure, because you may have run into an incompatibility between GNU C and traditional C (see Section 6.3 [Incompatibilities], page 131). These incompatibilities might be considered bugs, but they are inescapable consequences of valuable features.

  Or you may have a program whose behavior is undefined, which happened by chance to give the desired results with another C or C++ compiler.

  For example, in many nonoptimizing compilers, you can write '`x;`' at the end of a function instead of '`return x;`', with the same results. But the value of the function is undefined if `return` is omitted; it is not a bug when GCC produces different results.

  Problems often result from expressions with two increment operators, as in `f (*p++, *p++)`. Your previous compiler might have interpreted that expression the way you intended; GCC might interpret it another way. Neither compiler is wrong. The bug is in your code.

  After you have localized the error to a single source line, it should be easy to check for these things. If your program is correct and well defined, you have found a compiler bug.

- If the compiler produces an error message for valid input, that is a compiler bug.
- If the compiler does not produce an error message for invalid input, that is a compiler bug. However, you should note that your idea of "invalid input" might be my idea of "an extension" or "support for traditional practice".
- If you are an experienced user of one of the languages GCC supports, your suggestions for improvement of GCC are welcome in any case.

7.2 Where to Report Bugs

If you're using the compiler which came with MIPS SDE, send bug reports to `support@mips.com`.

There are GNU C bug mailing lists with much wider remits, but it would not usually be fair to snow them with reports from a piece of commercially-maintained software. But in some circumstances you may also be interested in the official GNU Compiler Collection list at '`gcc-bugs@gcc.gnu.org`'.

Please read '`<URL:http://www.gnu.org/software/gcc/bugs.html>`' for bug reporting instructions before you post a bug report.

Often people think of posting bug reports to the newsgroup instead of mailing them. This appears to work, but it has one problem which can be crucial: a newsgroup posting does not contain a mail path back to the sender. Thus, if maintainers need more information, they may be unable to reach you. For this reason, you should always send bug reports by mail to the proper mailing list.

As a last resort, send bug reports on paper to:

```
GNU Compiler Bugs
Free Software Foundation
59 Temple Place - Suite 330
Boston, MA 02111-1307, USA
```

7.3 How to Report Bugs

You may find additional and/or more up-to-date instructions at '`<URL:http://www.gnu.org/software/gcc/bugs.html>`'.

The fundamental principle of reporting bugs usefully is this: **report all the facts**. If you are not sure whether to state a fact or leave it out, state it!

Often people omit facts because they think they know what causes the problem and they conclude that some details don't matter. Thus, you might assume that the name of the variable you use in an example does not matter. Well, probably it doesn't, but one cannot be sure. Perhaps the bug is a stray memory reference which happens to fetch from the location where that name is stored in memory; perhaps, if the name were different, the contents of that location would fool the compiler into doing the right thing despite the bug. Play it safe and give a specific, complete example. That is the easiest thing for you to do, and the most helpful.

Keep in mind that the purpose of a bug report is to enable someone to fix the bug if it is not known. It isn't very important what happens if the bug is already known. Therefore, always write your bug reports on the assumption that the bug is not known.

Sometimes people give a few sketchy facts and ask, "Does this ring a bell?" This cannot help us fix a bug, so it is basically useless. We respond by asking for enough details to enable us to investigate. You might as well expedite matters by sending them to begin with.

Try to make your bug report self-contained. If we have to ask you for more information, it is best if you include all the previous information in your response, as well as the information that was missing.

Please report each bug in a separate message. This makes it easier for us to track which bugs have been fixed and to forward your bugs reports to the appropriate maintainer.

To enable someone to investigate the bug, you should include all these things:

- The version of GCC. You can get this by running it with the '`-v`' option.

  Without this, we won't know whether there is any point in looking for the bug in the current version of GCC.

- A complete input file that will reproduce the bug. If the bug is in the C preprocessor, send a source file and any header files that it requires. If the bug is in the compiler proper ('`cc1`'), send the preprocessor output generated by adding '`-save-temps`' to the compilation command (see Section 2.7 [Debugging Options], page 29). When you do this, use the same '`-I`', '`-D`' or '`-U`' options that you used in actual compilation. Then send the *input*.i or *input*.ii files generated.

  A single statement is not enough of an example. In order to compile it, it must be embedded in a complete file of compiler input; and the bug might depend on the details of how this is done.

  Without a real example one can compile, all anyone can do about your bug report is wish you luck. It would be futile to try to guess how to provoke the bug. For example, bugs in register allocation and reloading frequently depend on every little detail of the function they happen in.

  Even if the input file that fails comes from a GNU program, you should still send the complete test case. Don't ask the GCC maintainers to do the extra work of obtaining the program in question—they are all overworked as it is. Also, the problem may depend on what is in the header files on your system; it is unreliable for the GCC maintainers to try the problem with the header files available to them. By sending CPP output, you can eliminate this source of uncertainty and save us a certain percentage of wild goose chases.

- The command arguments you gave GCC to compile that example and observe the bug. For example, did you use '`-O`'? To guarantee you won't omit something important, list all the options.

  If we were to try to guess the arguments, we would probably guess wrong and then we would not encounter the bug.

- The type of machine you are using, and the operating system name and version number.

- The operands you gave to the `configure` command when you installed the compiler.

- A complete list of any modifications you have made to the compiler source. (We don't promise to investigate the bug unless it happens in an unmodified compiler. But if you've made modifications and don't tell us, then you are sending us on a wild goose chase.)

  Be precise about these changes. A description in English is not enough—send a context diff for them.

  Adding files of your own (such as a machine description for a machine we don't support) is a modification of the compiler source.

- Details of any other deviations from the standard procedure for installing GCC.

- A description of what behavior you observe that you believe is incorrect. For example, "The compiler gets a fatal signal," or, "The assembler instruction at line 208 in the output is incorrect."

Of course, if the bug is that the compiler gets a fatal signal, then one can't miss it. But if the bug is incorrect output, the maintainer might not notice unless it is glaringly wrong. None of us has time to study all the assembler code from a 50-line C program just on the chance that one instruction might be wrong. We need *you* to do this part!

Even if the problem you experience is a fatal signal, you should still say so explicitly. Suppose something strange is going on, such as, your copy of the compiler is out of synch, or you have encountered a bug in the C library on your system. (This has happened!) Your copy might crash and the copy here would not. If you *said* to expect a crash, then when the compiler here fails to crash, we would know that the bug was not happening. If you don't say to expect a crash, then we would not know whether the bug was happening. We would not be able to draw any conclusion from our observations.

If the problem is a diagnostic when compiling GCC with some other compiler, say whether it is a warning or an error.

Often the observed symptom is incorrect output when your program is run. Sad to say, this is not enough information unless the program is short and simple. None of us has time to study a large program to figure out how it would work if compiled correctly, much less which line of it was compiled wrong. So you will have to do that. Tell us which source line it is, and what incorrect result happens when that line is executed. A person who understands the program can find this as easily as finding a bug in the program itself.

- If you send examples of assembler code output from GCC, please use '`-g`' when you make them. The debugging information includes source line numbers which are essential for correlating the output with the input.
- If you wish to mention something in the GCC source, refer to it by context, not by line number.

  The line numbers in the development sources don't match those in your sources. Your line numbers would convey no useful information to the maintainers.
- Additional information from a debugger might enable someone to find a problem on a machine which he does not have available. However, you need to think when you collect this information if you want it to have any chance of being useful.

  For example, many people send just a backtrace, but that is never useful by itself. A simple backtrace with arguments conveys little about GCC because the compiler is largely data-driven; the same functions are called over and over for different RTL insns, doing different things depending on the details of the insn.

  Most of the arguments listed in the backtrace are useless because they are pointers to RTL list structure. The numeric values of the pointers, which the debugger prints in the backtrace, have no significance whatever; all that matters is the contents of the objects they point to (and most of the contents are other such pointers).

  In addition, most compiler passes consist of one or more loops that scan the RTL insn sequence. The most vital piece of information about such a loop—which insn it has reached—is usually in a local variable, not in an argument.

  What you need to provide in addition to a backtrace are the values of the local variables for several stack frames up. When a local variable or an argument is an RTX, first print its value and then use the GDB command `pr` to print the RTL expression that it points to. (If GDB doesn't run on your machine, use your debugger to call the function `debug_rtx` with the RTX as an argument.) In general, whenever a variable is a pointer, its value is no use without the data it points to.

Here are some things that are not necessary:

- A description of the envelope of the bug.

  Often people who encounter a bug spend a lot of time investigating which changes to the input file will make the bug go away and which changes will not affect it.

  This is often time consuming and not very useful, because the way we will find the bug is by running a single example under the debugger with breakpoints, not by pure deduction from a series of examples. You might as well save your time for something else.

  Of course, if you can find a simpler example to report *instead* of the original one, that is a convenience. Errors in the output will be easier to spot, running under the debugger will take less time, etc. Most GCC bugs involve just one function, so the most straightforward way to simplify an example is to delete all the function definitions except the one where the bug occurs. Those earlier in the file may be replaced by external declarations if the crucial function depends on them. (Exception: inline functions may affect compilation of functions defined later in the file.)

  However, simplification is not vital; if you don't want to do this, report the bug anyway and send the entire test case you used.

- In particular, some people insert conditionals '`#ifdef BUG`' around a statement which, if removed, makes the bug not happen. These are just clutter; we won't pay any attention to them anyway. Besides, you should send us cpp output, and that can't have conditionals.

- A patch for the bug.

  A patch for the bug is useful if it is a good one. But don't omit the necessary information, such as the test case, on the assumption that a patch is all we need. We might see problems with your patch and decide to fix the problem another way, or we might not understand it at all.

  Sometimes with a program as complicated as GCC it is very hard to construct an example that will make the program follow a certain path through the code. If you don't send the example, we won't be able to construct one, so we won't be able to verify that the bug is fixed.

  And if we can't understand what bug you are trying to fix, or why your patch should be an improvement, we won't install it. A test case will help us to understand.

  See Section 7.4 [Sending Patches], page 147, for guidelines on how to make it easy for us to understand and install your patches.

- A guess about what the bug is or what it depends on.

  Such guesses are usually wrong. Even I can't guess right about such things without first using the debugger to find the facts.

- A core dump file.

  We have no way of examining a core dump for your type of machine unless we have an identical system—and if we do have one, we should be able to reproduce the crash ourselves.

7.4 Sending Patches for GCC

If you would like to write bug fixes or improvements for the GNU C compiler, that is very helpful. Send suggested fixes to the patches mailing list, `gcc-patches@gcc.gnu.org`.

Please follow these guidelines so we can study your patches efficiently. If you don't follow these guidelines, your information might still be useful, but using it will take extra work. Maintaining

GNU C is a lot of work in the best of circumstances, and we can't keep up unless you do your best to help.

- Send an explanation with your changes of what problem they fix or what improvement they bring about. For a bug fix, just include a copy of the bug report, and explain why the change fixes the bug.

  (Referring to a bug report is not as good as including it, because then we will have to look it up, and we have probably already deleted it if we've already fixed the bug.)
- Always include a proper bug report for the problem you think you have fixed. We need to convince ourselves that the change is right before installing it. Even if it is right, we might have trouble judging it if we don't have a way to reproduce the problem.
- Include all the comments that are appropriate to help people reading the source in the future understand why this change was needed.
- Don't mix together changes made for different reasons. Send them *individually*.

  If you make two changes for separate reasons, then we might not want to install them both. We might want to install just one. If you send them all jumbled together in a single set of diffs, we have to do extra work to disentangle them—to figure out which parts of the change serve which purpose. If we don't have time for this, we might have to ignore your changes entirely.

  If you send each change as soon as you have written it, with its own explanation, then the two changes never get tangled up, and we can consider each one properly without any extra work to disentangle them.

  Ideally, each change you send should be impossible to subdivide into parts that we might want to consider separately, because each of its parts gets its motivation from the other parts.
- Send each change as soon as that change is finished. Sometimes people think they are helping us by accumulating many changes to send them all together. As explained above, this is absolutely the worst thing you could do.

  Since you should send each change separately, you might as well send it right away. That gives us the option of installing it immediately if it is important.
- Use '`diff -c`' to make your diffs. Diffs without context are hard for us to install reliably. More than that, they make it hard for us to study the diffs to decide whether we want to install them. Unidiff format is better than contextless diffs, but not as easy to read as '`-c`' format.

  If you have GNU diff, use '`diff -cp`', which shows the name of the function that each change occurs in.
- Write the change log entries for your changes. We get lots of changes, and we don't have time to do all the change log writing ourselves.

  Read the '`ChangeLog`' file to see what sorts of information to put in, and to learn the style that we use. The purpose of the change log is to show people where to find what was changed. So you need to be specific about what functions you changed; in large functions, it's often helpful to indicate where within the function the change was.

  On the other hand, once you have shown people where to find the change, you need not explain its purpose. Thus, if you add a new function, all you need to say about it is that it is new. If you feel that the purpose needs explaining, it probably does—but the explanation will be much more useful if you put it in comments in the code.

  If you would like your name to appear in the header line for who made the change, send us the header line.

- When you write the fix, keep in mind that we can't install a change that would break other systems.

  People often suggest fixing a problem by changing machine-independent files such as '`toplev.c`' to do something special that a particular system needs. Sometimes it is totally obvious that such changes would break GCC for almost all users. We can't possibly make a change like that. At best it might tell us how to write another patch that would solve the problem acceptably.

  Sometimes people send fixes that *might* be an improvement in general—but it is hard to be sure of this. It's hard to install such changes because we have to study them very carefully. Of course, a good explanation of the reasoning by which you concluded the change was correct can help convince us.

  The safest changes are changes to the configuration files for a particular machine. These are safe because they can't create new bugs on other machines.

  Please help us keep up with the workload by designing the patch in a form that is good to install.

8 How To Get Help with GCC

If you're a MIPS SDE customer on support, ask `support@mips.com`. Otherwise, if you need help installing, using or changing GNU CC, there are two ways to find it:

- Send a message to a suitable network mailing list. First try `gcc-bugs@gcc.gnu.org` or `bug-gcc@gnu.org`, and if that brings no response, try `gcc@gcc.gnu.org`.
- Look in the service directory for someone who might help you for a fee. The service directory is found in the file named '`SERVICE`' in the GCC distribution.

9 Contributing to GCC Development

If you would like to help pretest GCC releases to assure they work well, or if you would like to work on improving GCC, please contact the maintainers at `gcc@gcc.gnu.org`. A pretester should be willing to try to investigate bugs as well as report them.

If you'd like to work on improvements, please ask for suggested projects or suggest your own ideas. If you have already written an improvement, please tell us about it. If you have not yet started work, it is useful to contact `gcc@gcc.gnu.org` before you start; the maintainers may be able to suggest ways to make your extension fit in better with the rest of GCC and with other development plans.

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.

Linux and the GNU Project

Many computer users run a modified version of the GNU system every day, without realizing it. Through a peculiar turn of events, the version of GNU which is widely used today is more often known as “Linux”, and many users are not aware of the extent of its connection with the GNU Project.

There really is a Linux; it is a kernel, and these people are using it. But you can’t use a kernel by itself; a kernel is useful only as part of a whole system. The system in which Linux is typically used is a modified variant of the GNU system—in other words, a Linux-based GNU system.

Many users are not fully aware of the distinction between the kernel, which is Linux, and the whole system, which they also call “Linux”. The ambiguous use of the name doesn’t promote understanding.

Programmers generally know that Linux is a kernel. But since they have generally heard the whole system called “Linux” as well, they often envisage a history which fits that name. For example, many believe that once Linus Torvalds finished writing the kernel, his friends looked around for other free software, and for no particular reason most everything necessary to make a Unix-like system was already available.

What they found was no accident—it was the GNU system. The available free software added up to a complete system because the GNU Project had been working since 1984 to make one. The GNU Manifesto had set forth the goal of developing a free Unix-like system, called GNU. By the time Linux was written, the system was almost finished.

Most free software projects have the goal of developing a particular program for a particular job. For example, Linus Torvalds set out to write a Unix-like kernel (Linux); Donald Knuth set out to write a text formatter (TeX); Bob Scheifler set out to develop a window system (X Windows). It’s natural to measure the contribution of this kind of project by specific programs that came from the project.

If we tried to measure the GNU Project’s contribution in this way, what would we conclude? One CD-ROM vendor found that in their “Linux distribution”, GNU software was the largest single contingent, around 28% of the total source code, and this included some of the essential major components without which there could be no system. Linux itself was about 3%. So if you were going to pick a name for the system based on who wrote the programs in the system, the most appropriate single choice would be “GNU”.

But we don’t think that is the right way to consider the question. The GNU Project was not, is not, a project to develop specific software packages. It was not a project to develop a C compiler, although we did. It was not a project to develop a text editor, although we developed one. The GNU Project’s aim was to develop *a complete free Unix-like system*.

Many people have made major contributions to the free software in the system, and they all deserve credit. But the reason it is *a system*—and not just a collection of useful programs—is because the GNU Project set out to make it one. We wrote the programs that were needed to make a *complete* free system. We wrote essential but unexciting major components, such as the assembler and linker, because you can’t have a system without them. A complete system needs more than just programming tools, so we wrote other components as well, such as the Bourne Again SHell, the PostScript interpreter Ghostscript, and the GNU C library.

By the early 90s we had put together the whole system aside from the kernel (and we were also working on a kernel, the GNU Hurd, which runs on top of Mach). Developing this kernel has been a lot harder than we expected, and we are still working on finishing it.

Fortunately, you don't have to wait for it, because Linux is working now. When Linus Torvalds wrote Linux, he filled the last major gap. People could then put Linux together with the GNU system to make a complete free system: a Linux-based GNU system (or GNU/Linux system, for short).

Putting them together sounds simple, but it was not a trivial job. The GNU C library (called glibc for short) needed substantial changes. Integrating a complete system as a distribution that would work "out of the box" was a big job, too. It required addressing the issue of how to install and boot the system—a problem we had not tackled, because we hadn't yet reached that point. The people who developed the various system distributions made a substantial contribution.

The GNU Project supports GNU/Linux systems as well as *the* GNU system—even with funds. We funded the rewriting of the Linux-related extensions to the GNU C library, so that now they are well integrated, and the newest GNU/Linux systems use the current library release with no changes. We also funded an early stage of the development of Debian GNU/Linux.

We use Linux-based GNU systems today for most of our work, and we hope you use them too. But please don't confuse the public by using the name "Linux" ambiguously. Linux is the kernel, one of the essential major components of the system. The system as a whole is more or less the GNU system.

GNU GENERAL PUBLIC LICENSE

Version 2, June 1991

Preamble

The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software—to make sure the software is free for all its users. This General Public License applies to most of the Free Software Foundation's software and to any other program whose authors commit to using it. (Some other Free Software Foundation software is covered by the GNU Library General Public License instead.) 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 this service 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 make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the software, or if you modify it.

For example, if you distribute copies of such a program, whether gratis or for a fee, you must give the recipients all the rights that you have. 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.

We protect your rights with two steps: (1) copyright the software, and (2) offer you this license which gives you legal permission to copy, distribute and/or modify the software.

Also, for each author's protection and ours, we want to make certain that everyone understands that there is no warranty for this free software. If the software is modified by someone else and passed on, we want its recipients to know that what they have is not the original, so that any problems introduced by others will not reflect on the original authors' reputations.

Finally, any free program is threatened constantly by software patents. We wish to avoid the danger that redistributors of a free program will individually obtain patent licenses, in effect making the program proprietary. To prevent this, we have made it clear that any patent must be licensed for everyone's free use or not licensed at all.

The precise terms and conditions for copying, distribution and modification follow.

TERMS AND CONDITIONS FOR COPYING, DISTRIBUTION AND MODIFICATION

0. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The

"Program", below, refers to any such program or work, and a "work based on the Program" means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term "modification".) Each licensee is addressed as "you".

Activities other than copying, distribution and modification are not covered by this License; they are outside its scope. The act of running the Program is not restricted, and the output from the Program is covered only if its contents constitute a work based on the Program (independent of having been made by running the Program). Whether that is true depends on what the Program does.

1. You may copy and distribute 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 and disclaimer of warranty; keep intact all the notices that refer to this License and to the absence of any warranty; and give any other recipients of the Program a copy of this License along with the Program.

   You may charge a fee for the physical act of transferring a copy, and you may at your option offer warranty protection in exchange for a fee.

2. You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:

   a. You must cause the modified files to carry prominent notices stating that you changed the files and the date of any change.

   b. You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.

   c. If the modified program normally reads commands interactively when run, you must cause it, when started running for such interactive use in the most ordinary way, to print or display an announcement including an appropriate copyright notice and a notice that there is no warranty (or else, saying that you provide a warranty) and that users may redistribute the program under these conditions, and telling the user how to view a copy of this License. (Exception: if the Program itself is interactive but does not normally print such an announcement, your work based on the Program is not required to print an announcement.)

   These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.

   Thus, it is not the intent of this section to claim rights or contest your rights to work written entirely by you; rather, the intent is to exercise the right to control the distribution of derivative or collective works based on the Program.

In addition, mere aggregation of another work not based on the Program with the Program (or with a work based on the Program) on a volume of a storage or distribution medium does not bring the other work under the scope of this License.

3. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:

   a. Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

   b. Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

   c. Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

   The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable. However, as a special exception, the source code distributed need not include anything that is normally distributed (in either source or binary form) with the major components (compiler, kernel, and so on) of the operating system on which the executable runs, unless that component itself accompanies the executable.

   If distribution of executable or object code is made by offering access to copy from a designated place, then offering equivalent access to copy the source code from the same place counts as distribution of the source code, even though third parties are not compelled to copy the source along with the object code.

4. You may not copy, modify, sublicense, or distribute the Program except as expressly provided under this License. Any attempt otherwise to copy, modify, sublicense or distribute the Program is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

5. You are not required to accept this License, since you have not signed it. However, nothing else grants you permission to modify or distribute the Program or its derivative works. These actions are prohibited by law if you do not accept this License. Therefore, by modifying or distributing the Program (or any work based on the Program), you indicate your acceptance of this License to do so, and all its terms and conditions for copying, distributing or modifying the Program or works based on it.

6. Each time you redistribute the Program (or any work based on the Program), the recipient automatically receives a license from the original licensor to copy, distribute or modify the Program subject to these terms and conditions. You may not impose any further restrictions on the recipients' exercise of the rights granted herein. You are not responsible for enforcing compliance by third parties to this License.

7. If, as a consequence of a court judgment or allegation of patent infringement or for any other reason (not limited to patent issues), 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 distribute so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not distribute the Program at all. For example, if a patent license would not permit royalty-free redistribution of the Program by all those who receive copies directly or indirectly through you, then the only way you could satisfy both it and this License would be to refrain entirely from distribution of the Program.

   If any portion of this section is held invalid or unenforceable under any particular circumstance, the balance of the section is intended to apply and the section as a whole is intended to apply in other circumstances.

   It is not the purpose of this section to induce you to infringe any patents or other property right claims or to contest validity of any such claims; this section has the sole purpose of protecting the integrity of the free software distribution system, which is implemented by public license practices. Many people have made generous contributions to the wide range of software distributed through that system in reliance on consistent application of that system; it is up to the author/donor to decide if he or she is willing to distribute software through any other system and a licensee cannot impose that choice.

   This section is intended to make thoroughly clear what is believed to be a consequence of the rest of this License.

8. If the distribution and/or use of the Program is restricted in certain countries either by patents or by copyrighted interfaces, the original copyright holder who places the Program under this License may add an explicit geographical distribution limitation excluding those countries, so that distribution is permitted only in or among countries not thus excluded. In such case, this License incorporates the limitation as if written in the body of this License.

9. The Free Software Foundation may publish revised and/or new versions of the 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 a version number of this License which applies to it and “any later version”, you have the option of following the terms and conditions either of that version or of any later version published by the Free Software Foundation. If the Program does not specify a version number of this License, you may choose any version ever published by the Free Software Foundation.

10. If you wish to incorporate parts of the Program into other free programs whose distribution conditions are different, write to the author to ask for permission. For software which is copyrighted by the Free Software Foundation, write to the Free Software Foundation; we sometimes make exceptions for this. Our decision will be guided by the two goals of preserving the free status of all derivatives of our free software and of promoting the sharing and reuse of software generally.

NO WARRANTY

11. BECAUSE THE PROGRAM IS LICENSED FREE OF CHARGE, 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.

12. IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MAY MODIFY AND/OR REDISTRIBUTE 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.

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 convey 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) *yyyy 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 2 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, write to the Free Software
> Foundation, Inc., 59 Temple Place - Suite 330, Boston, MA 02111-1307, USA.

Also add information on how to contact you by electronic and paper mail.

If the program is interactive, make it output a short notice like this when it starts in an interactive mode:

> Gnomovision version 69, Copyright (C) *yyyy name of author*
> Gnomovision 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, the commands you use may be called something other than 'show w' and 'show c'; they could even be mouse-clicks or menu items—whatever suits your program.

You should also get your employer (if you work as a programmer) or your school, if any, to sign a "copyright disclaimer" for the program, if necessary. Here is a sample; alter the names:

> Yoyodyne, Inc., hereby disclaims all copyright interest in the program
> 'Gnomovision' (which makes passes at compilers) written by James Hacker.
>
> *signature of Ty Coon*, 1 April 1989
> Ty Coon, President of Vice

This 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 Library General Public License instead of this License.

Contributors to GCC

In addition to Richard Stallman, several people have written parts of GCC.

- The idea of using RTL and some of the optimization ideas came from the program PO written at the University of Arizona by Jack Davidson and Christopher Fraser. See "Register Allocation and Exhaustive Peephole Optimization", Software Practice and Experience 14 (9), Sept. 1984, 857-866.
- Paul Rubin wrote most of the preprocessor.
- Leonard Tower wrote parts of the parser, RTL generator, and RTL definitions, and of the Vax machine description.
- Ted Lemon wrote parts of the RTL reader and printer.
- Jim Wilson implemented loop strength reduction and some other loop optimizations.
- Nobuyuki Hikichi of Software Research Associates, Tokyo, contributed the support for the Sony NEWS machine.
- Charles LaBrec contributed the support for the Integrated Solutions 68020 system.
- Michael Tiemann of Cygnus Support wrote the front end for C++, as well as the support for inline functions and instruction scheduling. Also the descriptions of the National Semiconductor 32000 series cpu, the SPARC cpu and part of the Motorola 88000 cpu.
- Gerald Baumgartner added the signature extension to the C++ front-end.
- Jan Stein of the Chalmers Computer Society provided support for Genix, as well as part of the 32000 machine description.
- Randy Smith finished the Sun FPA support.
- Robert Brown implemented the support for Encore 32000 systems.
- David Kashtan of SRI adapted GCC to VMS.
- Alex Crain provided changes for the 3b1.
- Greg Satz and Chris Hanson assisted in making GCC work on HP-UX for the 9000 series 300.
- William Schelter did most of the work on the Intel 80386 support.
- Christopher Smith did the port for Convex machines.
- Paul Petersen wrote the machine description for the Alliant FX/8.
- Dario Dariol contributed the four varieties of sample programs that print a copy of their source.
- Alain Lichnewsky ported GCC to the Mips cpu.
- Devon Bowen, Dale Wiles and Kevin Zachmann ported GCC to the Tahoe.
- Jonathan Stone wrote the machine description for the Pyramid computer.
- Gary Miller ported GCC to Charles River Data Systems machines.
- 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.
- Richard Kenner and Michael Tiemann jointly developed reorg.c, the delay slot scheduler.

- Mike Meissner and Tom Wood of Data General finished the port to the Motorola 88000.
- Masanobu Yuhara of Fujitsu Laboratories implemented the machine description for the Tron architecture (specifically, the Gmicro).
- NeXT, Inc. donated the front end that supports the Objective C language.
- James van Artsdalen wrote the code that makes efficient use of the Intel 80387 register stack.
- Mike Meissner at the Open Software Foundation finished the port to the MIPS cpu, including adding ECOFF debug support, and worked on the Intel port for the Intel 80386 cpu. Later at Cygnus Support, he worked on the rs6000 and PowerPC ports.
- 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.
- Torbjorn Granlund implemented multiply- and divide-by-constant optimization, improved long long support, and improved leaf function register allocation.
- Mike Stump implemented the support for Elxsi 64 bit CPU.
- John Wehle added the machine description for the Western Electric 32000 processor used in several 3b series machines (no relation to the National Semiconductor 32000 processor).
- Holger Teutsch provided the support for the Clipper cpu.
- Kresten Krab Thorup wrote the run time support for the Objective C language.
- Stephen Moshier contributed the floating point emulator that assists in cross-compilation and permits support for floating point numbers wider than 64 bits.
- David Edelsohn contributed the changes to RS/6000 port to make it support the PowerPC and POWER2 architectures.
- Steve Chamberlain wrote the support for the Hitachi SH processor.
- Peter Schauer wrote the code to allow debugging to work on the Alpha.
- Oliver M. Kellogg of Deutsche Aerospace contributed the port to the MIL-STD-1750A.
- Michael K. Gschwind contributed the port to the PDP-11.
- David Reese of Sun Microsystems contributed to the Solaris on PowerPC port.

Index

B

C

D

E

F

G

H

I

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Z

Short Contents

Table of Contents