IEEE P1003.2 Draft 11.2 – September 1991

Copyright © 1991 by the
Institute of Electrical and Electronics Engineers, Inc.
345 East 47th Street
New York, NY 10017, USA
All rights reserved as an unpublished work.

This is an unapproved and unpublished IEEE Standards Draft, subject to change. The publication, distribution, or copying of this draft, as well as all derivative works based on this draft, is expressly prohibited except as set forth below.

Permission is hereby granted for IEEE Standards Committee participants to reproduce this document for purposes of IEEE standardization activities only, and subject to the restrictions contained herein.

Permission is hereby also granted for member bodies and technical committees of ISO and IEC to reproduce this document for purposes of developing a national position, subject to the restrictions contained herein.

Permission is hereby also granted to the preceding entities to make limited copies of this document in an electronic form only for the stated activities.

The following restrictions apply to reproducing or transmitting the document in any form: 1) all copies or portions thereof must identify the document's IEEE project number and draft number, and must be accompanied by this entire notice in a prominent location; 2) no portion of this document may be redistributed in any modified or abridged form without the prior approval of the IEEE Standards Department.

Other entities seeking permission to reproduce this document, or any portion thereof, for standardization or other activities, must contact the IEEE Standards Department for the appropriate license.

Use of information contained in this unapproved draft is at your own risk.

IEEE Standards Department
Copyright and Permissions
445 Hoes Lane, P.O. Box 1331
Piscataway, NJ 08855-1331, USA
+1 (908) 562-3800
+1 (908) 562-1571 [FAX]

4.24 `find` — Find files

4.24.1 Synopsis

`find` *path* ... [*operand_expression* ...]

4.24.2 Description

The `find` utility shall recursively descend the directory hierarchy from each file specified by *path*, evaluating a Boolean expression composed of the primaries described in 4.24.4 for each file encountered.

The `find` utility shall be able to descend to arbitrary depths in a file hierarchy and shall not fail due to path length limitations (unless a path operand specified by the application exceeds {PATH_MAX} requirements).

The `find` utility requires that the underlying system provides information equivalent to the *st_dev*, *st_mode*, *st_nlink*, *st_uid*, *st_gid*, *st_size*, *st_atime*, *st_mtime*, and *st_ctime* members of *struct stat* described by POSIX.1 {8} 5.6 and conforming to the *file times update* definition in 2.2.2.69.

4.24.3 Options

None.

4.24.4 Operands

The following operands shall be supported by the implementation:

The *path* operand is a pathname of a starting point in the directory hierarchy.

The first argument that starts with a `-`, or is a `!` or a `(`, and all subsequent arguments shall be interpreted as an *expression* made up of the following primaries and operators. In the descriptions, wherever *n* is used as a primary argument, it shall be interpreted as a decimal integer optionally preceded by a plus (+) or minus (–) sign, as follows:

`+`*n* More than *n*

n Exactly *n*

`-`*n* Less than *n*

Implementations shall recognize the following primaries: *Editor's Note: These primaries have been sorted alphabetically, without diff marks.*

`–atime` *n* The primary shall evaluate as true if the file access time subtracted from the initialization time is *n*–1 to *n* multiples of 24 hours. The initialization time shall be a time between the invocation of the `find` utility and the first

Copyright © 1991 IEEE. All rights reserved.
This is an unapproved IEEE Standards Draft, subject to change.

5943 access by that invocation of the `find` utility to any file
5944 specified by its *path* operands.

5945 `-ctime` *n* The primary shall evaluate as true if the time of last
5946 change of file status information subtracted from the ini-
5947 tialization time is *n*−1 to *n* multiples of 24 hours. The ini-
5948 tialization time shall be a time between the invocation of
5949 the `find` utility and the first access by that invocation of
5950 the `find` utility to any file specified by its *path* operands.

5951 `-depth` The primary always shall evaluate as true; it shall cause
5952 descent of the directory hierarchy to be done so that all
5953 entries in a directory are acted on before the directory
5954 itself. If a `-depth` primary is not specified, all entries in a
5955 directory shall be acted on after the directory itself. If any
5956 `-depth` primary is specified, it shall apply to the entire
5957 expression even if the `-depth` primary would not nor-
5958 mally be evaluated.

5959 `-exec` *utility_name* `[`*argument* ...`];`
5960 The primary shall evaluate as true if the invoked utility
5961 *utility_name* returns a zero value as exit status. The end
5962 of the primary expression shall be punctuated by a semi-
5963 colon. A *utility_name* or *argument* containing only the
5964 two characters `{}` shall be replaced by the current path-
5965 name. If a utility_name or argument string contains the
5966 two characters `{}`, but not just the two characters `{}`, it is
5967 implementation defined whether `find` replaces those two
5968 characters with the current pathname or uses the string
5969 without change. The current directory for the invocation
5970 of *utility_name* shall be the same as the current directory
5971 when the `find` utility was started. If the *utility_name*
5972 names any of the special built-in utilities in 3.14, the
5973 results are undefined.

5974 `-group` *gname* The primary shall evaluate as true if the file belongs to
5975 the group *gname* . If *gname* is a decimal integer and the
5976 *getgrnam* () (or equivalent) function does not return a
5977 valid group name, *gname* shall be interpreted as a group
5978 ID.

5979 `-links` *n* The primary shall evaluate as true if the file has *n* links.

5980 `-mtime` *n* The primary shall evaluate as true if the file modification
5981 time subtracted from the initialization time is *n*−1 to *n*
5982 multiples of 24 hours. The initialization time shall be a
5983 time between the invocation of the `find` utility and the
5984 first access by that invocation of the `find` utility to any
5985 file specified by its *path* operands.

Copyright © 1991 IEEE. All rights reserved.
This is an unapproved IEEE Standards Draft, subject to change.

5986 `–name` *pattern* The primary shall evaluate as true if the basename of the
5987 filename being examined matches *pattern* using the pat-
5988 tern matching notation described in 3.13.

5989 `–newer` *file* The primary shall evaluate as true if the modification
5990 time of the current file is more recent than the
5991 modification time of the file named by the pathname *file*.

5992 `–nogroup` The primary shall evaluate as true if the file belongs to a
5993 group ID for which the POSIX.1 {8} *getgrgid* () (or
5994 equivalent) function returns **NULL**.

5995 `–nouser` The primary shall evaluate as true if the file belongs to a
5996 user ID for which the POSIX.1 {8} *getpwuid* () (or
5997 equivalent) function returns **NULL**.

5998 `-ok` *utility_name* [*argument* ...] ;
5999 The `–ok` primary shall be equivalent to `–exec`, except that
6000 `find` shall request affirmation of the invocation of
6001 *utility_name* using the current file as an argument by
6002 writing to standard error as, described in 4.24.6.2. If the
6003 response on standard input is affirmative, the utility shall
6004 be invoked. Otherwise, the command shall not be invoked
6005 and the value of the `–ok` operand shall be false.

6006 `-perm [-]`*mode* The *mode* argument is used to represent file mode bits. It
6007 shall be identical in format to the *symbolic_mode* operand
6008 described in 4.7, and shall be interpreted as follows. To
6009 start, a template shall be assumed with all file mode bits
6010 cleared. An *op* symbol of `+` shall set the appropriate mode
6011 bits in the template; `-` shall clear the appropriate bits; `=`
6012 shall set the appropriate mode bits, without regard to the
6013 contents of process's file mode creation mask. The *op* sym-
6014 bol of `-` cannot be the first character of *mode*.

6015 If the hyphen is omitted, the primary shall evaluate as
6016 true when the file permission bits exactly match the value
6017 of the resulting template.

6018 Otherwise, if *mode* is prefixed by a hyphen, the primary
6019 shall evaluate as true if at least all the bits in the result-
6020 ing template are set in the file permission bits.

6021 `-perm [-]`*onum* (Obsolescent.) If the hyphen is omitted, the primary shall
6022 evaluate as true when the file permission bits exactly
6023 match the value of the octal number *onum* and only the
6024 bits corresponding to the octal mask 07777 shall be com-
6025 pared. (See the description of the octal *mode* in 4.7.) Oth-
6026 erwise, if *onum* is prefixed by a hyphen, the primary shall
6027 evaluate as true if at least all of the bits specified in *onum*
6028 that are also set it the octal mask 07777 are set.

Copyright © 1991 IEEE. All rights reserved.
This is an unapproved IEEE Standards Draft, subject to change.

6029 **−print** The primary always shall evaluate as true; it shall cause
6030 the current pathname to be written to standard output.

6031 **−prune** The primary always shall evaluate as true; it shall cause
6032 **find** not to descend the current pathname if it is a direc-
6033 tory. If the **−depth** primary is specified, the **−prune** pri-
6034 mary shall have no effect.

6035 **-size** *n*[**c**] The primary shall evaluate as true if the file size in bytes,
6036 divided by 512 and rounded up to the next integer, is *n*. If
6037 *n* is followed by the character **c**, the size shall be in bytes.

6038 **−type** *c* The primary shall evaluate as true if the type of the file is
6039 *c*, where *c* is **b**, **c**, **d**, **p**, or **f** for block special file, character
6040 special file, directory, FIFO , or regular file, respectively.

6041 **−user** *uname* The primary shall evaluate as true if the file belongs to
6042 the user *uname*. If *uname* is a decimal integer and the
6043 *getpwnam* () (or equivalent) function does not return a
6044 valid user name, *uname* shall be interpreted as a user ID.

6045 **−xdev** The primary always shall evaluate as true; it shall cause
6046 **find** not to continue descending past directories that
6047 have a different device ID (*st_dev* , see POSIX .1 {8} 5.6.2).
6048 If any **−xdev** primary is specified, it shall apply to the
6049 entire expression even if the **−xdev** primary would not
6050 normally be evaluated.

6051 The primaries can be combined using the following operators (in order of decreas-
6052 ing precedence):

6053 **(** *expression* **)** True if *expression* is true.

6054 **!** *expression* Negation of a primary; the unary NOT operator.

6055 *expression* [**-a**] *expression*
6056 Conjunction of primaries; the AND operator shall be
6057 implied by the juxtaposition of two primaries or made
6058 explicit by the optional **−a** operator. The second expres-
6059 sion shall not be evaluated if the first expression is false.

6060 *expression* **−o** *expression*
6061 Alternation of primaries; the OR operator. The second
6062 expression shall not be evaluated if the first expression is
6063 true.

6064 If no *expression* is present, **−print** shall be used as the expression. Otherwise, if
6065 the given expression does not contain any of the primaries **−exec**, **−ok**, or
6066 **−print**, the given expression shall be effectively replaced by:

6067 **(** *given_expression* **) -print**

6068 The **−user**, **−group**, and **−newer** primaries each shall evaluate their respective
6069 arguments only once.

Copyright © 1991 IEEE. All rights reserved.
This is an unapproved IEEE Standards Draft, subject to change.

6070 ### 4.24.5 External Influences

6071 #### 4.24.5.1 Standard Input

6072 If the –ok primary is used, the response shall be read from the standard input.
6073 An entire line shall be read as the response. Otherwise, the standard input shall
6074 not be used.

6075 #### 4.24.5.2 Input Files

6076 None.

6077 #### 4.24.5.3 Environment Variables

6078 The following environment variables shall affect the execution of find:

6079 **LANG** This variable shall determine the locale to use for the
6080 locale categories when both **LC_ALL** and the correspond-
6081 ing environment variable (beginning with **LC_**) do not
6082 specify a locale. See 2.6.

6083 **LC_ALL** This variable shall determine the locale to be used to over-
6084 ride any values for locale categories specified by the set-
6085 tings of **LANG** or any environment variables beginning
6086 with **LC_**.

6087 **LC_COLLATE** This variable shall determine the locale for the behavior of
6088 ranges, equivalence classes, and multicharacter collating
6089 elements used in the pattern matching notation for the
6090 –name option and in the extended regular expression
6091 defined for the yesexpr locale keyword in the
6092 LC_MESSAGES category.

6093 **LC_CTYPE** This variable shall determine the locale for the interpreta-
6094 tion of sequences of bytes of text data as characters (e.g.,
6095 single- versus multibyte characters in arguments), the
6096 behavior of character classes within the pattern matching
6097 notation used for the –name option, and the behavior of
6098 character classes within regular expressions used in the
6099 extended regular expression defined for the yesexpr
6100 locale keyword in the LC_MESSAGES category.

6101 **LC_MESSAGES** This variable shall determine the processing of affirmative
6102 responses and the language in which messages should be
6103 written.

6104 **PATH** This variable shall determine the location of the
6105 *utility_name* for the –exec and –ok primaries, as
6106 described in 2.6.

Copyright © 1991 IEEE. All rights reserved.
This is an unapproved IEEE Standards Draft, subject to change.

6107 #### 4.24.5.4 Asynchronous Events

6108 Default.

6109 ### 4.24.6 External Effects

6110 #### 4.24.6.1 Standard Output

6111 The `-print` primary shall cause the current pathnames to be written to standard
6112 output. The format shall be:

6113 `"%s\n"`, *<path>*

6114 #### 4.24.6.2 Standard Error

6115 The `-ok` primary shall write a prompt to standard error containing at least the
6116 utility_name to be invoked and the current pathname. In the POSIX Locale, the
6117 last non-`<blank>` character in the prompt shall be `?`. The exact format used is
6118 unspecified.

6119 Otherwise, the standard error shall be used only for diagnostic messages.

6120 #### 4.24.6.3 Output Files

6121 None.

6122 ### 4.24.7 Extended Description

6123 None.

6124 ### 4.24.8 Exit Status

6125 The `find` utility shall exit with one of the following values:

6126 0 All *path* operands were traversed successfully.

6127 >0 An error occurred.

6128 ### 4.24.9 Consequences of Errors

6129 Default.

Copyright © 1991 IEEE. All rights reserved.
This is an unapproved IEEE Standards Draft, subject to change.

4.24.10 Rationale. *(This subclause is not a part of P1003.2)*

Examples, Usage

When used in operands, pattern matching notation, semicolons, opening parentheses, and closing parentheses are special to the shell and must be quoted (see 3.2).

The following command:

```
find / \( -name tmp -o -name '*.xx' \) \
        -atime +7 -exec rm {} \;
```

removes all files named `tmp` or ending in `.xx` that have not been accessed for seven or more 24-hour periods.

The following command:

```
find . -perm -o+w,+s
```

prints (`-print` is assumed) the names of all files in or below the current directory, with all of the file permission bits S_ISUID , S_ISGID , and S_IWOTH set.

The `–prune` primary was adopted from later releases of 4.3 BSD and the third edition of the *SVID* . The following command recursively prints pathnames of all files in the current directory and below, but skips directories named SCCS and files in them.

```
find . -name SCCS -prune -o -print
```

The following command behaves as in the previous example, but prints the names of the SCCS directories.

```
find . -print -name SCCS -prune
```

The following command is roughly equivalent to the `–nt` extension to `test`:

```
if [ -n "$(find file1 -prune -newer file2)" ]; then
        printf %s\\n "file1 is newer than file2"
fi
```

History of Decisions Made

The historical `–a` operator is kept as an optional operator for compatibility with existing shell scripts even though it is redundant with expression concatenation.

The symbolic means of specifying file permission bits, based on `chmod`, was added in response to numerous balloting objections that `find` was the only remaining utility to not support this method. The warning about a leading *Op* of `-` is to avoid ambiguity with the optional leading hyphen. Since the initial mode is all bits off, there are not any symbolic modes that need to use `-` as the first character. The bit that is traditionally used for sticky (historically 01000) is still specified in the `–perm` primary using the octal number argument form. Since this bit is not defined by POSIX .1 {8} or POSIX. 2, applications must not assume that it actually refers to the traditional sticky bit.

Copyright © 1991 IEEE. All rights reserved.
This is an unapproved IEEE Standards Draft, subject to change.

The descriptions of how the - modifier on the *mode* and *onum* arguments to the –perm primary affects processing has been documented here to match the way it behaves in practice on historical BSD and System V implementations. System V and BSD documentation both describe it in terms of checking additional bits; in fact, it uses the same bits, but checks for having at least all of the matching bits set instead of having exactly the matching bits set.

The exact format of the interactive prompts is unspecified. Only the general nature of the contents of prompts are specified, because:

(1) Implementations may desire more descriptive prompts than those used on historical implementations.

(2) Since the traditional prompt strings do not terminate with <newline>s, there is no portable way for another program to interact with the prompts of this utility via pipes.

Therefore, an application using this prompting option relies on the system to provide the most suitable dialogue directly with the user, based on the general guidelines specified.

The –name *file* operand was changed to use the shell pattern matching notation so that find is consistent with other utilities using pattern matching.

For the –type *c* operand, implementors of symbolic links should consider l (the letter ell) for symbolic links. Implementations that support sockets also use –type s for sockets. Implementations planning to add options to allow find to follow symbolic links or treat them as special files, should consider using –follow as used in BSD and System V Release 4 as a guide.

The –size operand refers to the size of a file, rather than the number of blocks it may occupy in the file system. The intent is that the POSIX .1 {8} *st_size* field should be used, not the *st_blocks* found in historical implementations. There are at least two reasons for this:

— In both System V and BSD, find only uses *st_size* in size calculations for the operands specified by POSIX. 2. (BSD uses *st_blocks* only when processing the –ls primary.)

— Users will usually be thinking of size in terms of the size of the file in bytes, which is also used by the ls utility for the output from the –l option. (In both System V in BSD, ls uses *st_size* for the –l option size field and uses *st_blocks* for the ls –s calculations. POSIX. 2 does not specify ls –s.)

The descriptions of –atime, –ctime, and –mtime were changed from the *SVID* 's description of *n* ''days'' to ''24-hour periods.'' For example, a file accessed at 23:59 will be selected by

```
find . -atime -1 -print
```

at 00:01 the next day (less than 24 hours later, not more than one day ago); the midnight boundary between days has no effect on the 24-hour calculation. The description is also different in terms of the exact timeframe for the *n* case (versus the +*n* or –*n*), but it matches all known historical implementations. It refers to

Copyright © 1991 IEEE. All rights reserved.
This is an unapproved IEEE Standards Draft, subject to change.

one 24-hour period in the past, not any time from the beginning of that period to the current time. For example, **–atime 3** is true if the file was accessed any time in the period from 72 to 48 hours ago.

Historical implementations do not modify **{}** when it appears as a substring of an **–exec** or **–ok** *utility_name* or argument string. There have been numerous user requests for this extension, so this standard allows the desired behavior. At least one recent implementation does support this feature, but ran into several problems in managing memory allocation and dealing with multiple occurrences of **{}** in a string while it was being developed, so it is not yet required behavior.

Assuming the presence of **–print** was added at the request of several working group members to correct a historical pitfall that plagues novice users. It is entirely upward compatible from the historical System V **find** utility and should be easy to implement. In its simplest form (**find** *directory*), it could be confused with the historical BSD fast **find**. The BSD developers agree that adding **–print** as a default expression is the right thing to do and believe that the fast **find** functionality should have been/should be provided by a separate utility. They suggest that the new utility be called **locate**.

4.25 fold — Fold lines

4.25.1 Synopsis

fold [–bs][–w *width* **][***file* **. . .]**

4.25.2 Description

The **fold** utility is a filter that shall fold lines from its input files, breaking the lines to have a maximum of *width* column positions (or bytes, if the **–b** option is specified). Lines shall be broken by the insertion of a **<newline>** character such that each output line (referred to later in this clause as a segment) is the maximum width possible that does not exceed the specified number of column positions (or bytes). A line shall not be broken in the middle of a character. The behavior is undefined if *width* is less than the number of columns any single character in the input would occupy.

If the **<carriage-return>**, **<backspace>**, or **<tab>** characters are encountered in the input, and the **–b** option is not specified, they shall be treated specially:

<carriage-return>
: The current count of line width shall be set to zero. The **fold** utility shall not insert a **<newline>** immediately before or after any **<carriage-return>**.

<backspace>
: The current count of line width shall be decremented by one, although the count never shall become negative. The **fold** utility shall not

Copyright © 1991 IEEE. All rights reserved.
This is an unapproved IEEE Standards Draft, subject to change.