NAME

regcmp – regular expression compile

SYNOPSIS

regcmp [–] *filename*. . .

DESCRIPTION

The **regcmp** command performs a function similar to **regcmp** and, in most cases, precludes the need for calling **regcmp** from C programs. Bypassing **regcmp** saves on both execution time and program size. The command **regcmp** compiles the regular expressions in *filename* and places the output in *filename***.i**.

OPTIONS

– If the – option is used, the output is placed in *filename***.c**. The format of entries in *filename* is a name (C variable) followed by one or more blanks followed by one or more regular expressions enclosed in double quotes. The output of **regcmp** is C source code. Compiled regular expressions are represented as **extern char** vectors. *filename***.i** files may thus be **#include**d in C programs, or *filename***.c** files may be compiled and later loaded. In the C program that uses the **regcmp** output, **regex(abc,line)** applies the regular expression named **abc** to **line**. Diagnostics are self-explanatory.

EXAMPLES

```
name   "([A–Za–z][A–Za–z0–9_]*)$0"
telno  " \({0,1}([2–9][01][1–9])$0\){0,1} *"
       "([2–9][0–9]{2})$1[ –]{0,1}"
       "([0–9]{4})$2"
```

The three arguments to **telno** shown above must all be entered on one line.

In the C program that uses the **regcmp** output,

regex(telno, line, area, exch, rest)

applies the regular expression named **telno** to **line**.

ENVIRONMENT

If any of the **LC_*** variables (**LC_CTYPE, LC_MESSAGES, LC_TIME, LC_COLLATE, LC_NUMERIC,** and **LC_MONETARY**) (see **environ**(5)) are not set in the environment, the operational behavior of **regcmp** for each corresponding locale category is determined by the value of the **LANG** environment variable. If **LC_ALL** is set, its contents are used to override both the **LANG** and the other **LC_*** variables. If none of the above variables is set in the environment, the "C" (U.S. style) locale determines how **regcmp** behaves.

LC_CTYPE

Determines how **regcmp** handles characters. When **LC_CTYPE** is set to a valid value, **regcmp** can display and handle text and filenames containing valid characters for that locale. **regcmp** can display and handle Extended Unix Code (EUC) characters where any individual character can be 1, 2, or 3 bytes wide. **regcmp** can also handle EUC characters of 1, 2, or more column widths. In the "C" locale, only characters from ISO 8859-1 are valid.

LC_MESSAGES
Determines how diagnostic and informative messages are presented. This includes the language and style of the messages, and the correct form of affirmative and negative responses. In the "C" locale, the messages are presented in the default form found in the program itself (in most cases, U.S. English).

SEE ALSO

regcmp(3G), **environ**(5)