

NAME

PCRE - Perl-compatible regular expressions

SIZE AND OTHER LIMITATIONS

There are some size limitations in PCRE but it is hoped that they will never in practice be relevant.

The maximum length of a compiled pattern is approximately 64K data units (bytes for the 8-bit library, 16-bit units for the 16-bit library, and 32-bit units for the 32-bit library) if PCRE is compiled with the default internal linkage size, which is 2 bytes for the 8-bit and 16-bit libraries, and 4 bytes for the 32-bit library. If you want to process regular expressions that are truly enormous, you can compile PCRE with an internal linkage size of 3 or 4 (when building the 16-bit or 32-bit library, 3 is rounded up to 4). See the **README** file in the source distribution and the **pcrebuild** documentation for details. In these cases the limit is substantially larger. However, the speed of execution is slower.

All values in repeating quantifiers must be less than 65536.

There is no limit to the number of parenthesized subpatterns, but there can be no more than 65535 capturing subpatterns. There is, however, a limit to the depth of nesting of parenthesized subpatterns of all kinds. This is imposed in order to limit the amount of system stack used at compile time. The limit can be specified when PCRE is built; the default is 250.

There is a limit to the number of forward references to subsequent subpatterns of around 200,000. Repeated forward references with fixed upper limits, for example, `(?2){0,100}` when subpattern number 2 is to the right, are included in the count. There is no limit to the number of backward references.

The maximum length of name for a named subpattern is 32 characters, and the maximum number of named subpatterns is 10000.

The maximum length of a name in a `(*MARK)`, `(*PRUNE)`, `(*SKIP)`, or `(*THEN)` verb is 255 for the 8-bit library and 65535 for the 16-bit and 32-bit libraries.

The maximum length of a subject string is the largest positive number that an integer variable can hold. However, when using the traditional matching function, PCRE uses recursion to handle subpatterns and indefinite repetition. This means that the available stack space may limit the size of a subject string that can be processed by certain patterns. For a discussion of stack issues, see the **prestack** documentation.

AUTHOR

Philip Hazel

University Computing Service
Cambridge CB2 3QH, England.

REVISION

Last updated: 05 November 2013

Copyright (c) 1997-2013 University of Cambridge.