Regular Expressions

Paolo Bettelini

Contents

1 Operators

1.1 Anchors

Symbol	Description
^	Start of string or line
$	End of string or line
\A	Start of string
\Z	End of string
\b	Word boundary
\B	Not word boundary
\<	Start of word
\>	End of word

1.2 Quantifiers

Symbol	Description
*	Match 0 or more times
+	Match 1 or more times
?	Match 0 or 1 time
{n}	Match n times
{n,}	Match n or more times
{n,m}	Match from n to m times
?	You can add a ? after a quantifier to make it *un-greedy*. The engine will match as few characters as possible.

1.3 Escaping

Symbol	Description
\	Escape character
\Q	Start literal sequence
\E	End literal sequence

1.4 Ranges

Symbol	Description
[· · ·]	A range. E.g. [abcz], [0-9], [0-9A-Z], [a-f].
[^· · ·]	Not range

1.5 Characters

Symbol	Description
\c	Control character
\s	White space
\S	Not white space
\d	Digit
\D	Not digit
\w	Word
\W	Not word
\x	Hexadecimal digit
\O	Octal digit

1.6 POSIX

Symbol	Description
[:upper:]	Upper case letters
[:lower:]	Lower case letters
[:alpha:]	All letters
[:alnum:]	Digits and letters
[:digit:]	Digits
[:xdigit:]	Hexadecimal digits
[:punct:]	Punctuation
[:blank:]	Space and tab
[:blank:]	Blank characters
[:cntrl:]	Control characters
[:graph:]	Printed characters
[:print:]	Printed characters and spaces
[:word:]	Digits, letters and underscore

1.7 Special characters

Symbol	Description
\n	New line
\r	Carriage return
\t	Tab
\v	Vertical tab
\f	Form feed
\XXX	Octal character *XXX*
\xHH	Hex character *HH*

1.8 String Replacement

Symbol	Description
$n	Nth non-passive group
$2	**xyz** in **/^(abc(xyz))$/**
$1	**xyz** in **/^(?:abc(xyz))$/**
$‘	Before matched string
$’	After matches string
$+	Last matches string
$&	Entire matched string

1.9 Assertions

Symbol	Description
?=	Lookahead assertion
?!	Negative lookahead
?<=	Lookbehind assertion
?!=	Negative lookbehind
?<!	Negative lookbehind
?>	Once-only Subexpression
?()	Condition [if then]
?()|	Condition [if then else]
?#	Comment

1.10 Pattern modifiers

Symbol	Description
g	Global match
i *	Case-insensitive
m *	Multiple lines
s *	Treat string as single line
x *	Allow comments and whitespace in pattern
e *	Evaluate replacement
U *	Un-greedy pattern

1.11 Other

Symbol	Description
(...)	Group
?:	Non-capturing group
|	Or operator