

Generic Notation for Regular Expressions

Many programming tools employ regular expressions, from utilities such as `grep` to programming languages like Perl. The following is a summary of the most common regular expression constructs.

<i>Notation</i>	<i>Description</i>
x	Any character x , except: () * + ? [] . ^ \ \$
$\backslash x$	Any character x
AB	A followed by B
$A B$	Choose one of A and B
(A)	A as a capturing group
A^*	A zero or more times
A^+	A one or more times
$A?$	A one or zero times
$A^*?$	A zero or more times, reluctant version
$A^+?$	A one or more times, reluctant version
$A??$	A one or zero times, reluctant version
$A\{n\}$	A exactly n times
$A\{n,\}$	A at least n times
$A\{n,m\}$	A at least n times, but no more than m times
$[abc]$	Chose one of the characters a , b , or c .
$[^abc]$	Chose one character except a , b , or c .
$[a-z]$	Chose one character from a to z inclusively.
$\backslash w$	Word character. Same as: $[a-zA-Z0-9_]$
$\backslash W$	Non-word character. Same as: $[^a-zA-Z0-9_]$
$\backslash d$	Digit character. Same as: $[0-9]$
$\backslash D$	Non-digit character. Same as: $[^0-9]$
$\backslash s$	Space character. Same as: $[\backslash t\backslash n\backslash f\backslash r]$
$\backslash S$	Non-space character. Same as: $[^\backslash t\backslash n\backslash f\backslash r]$
$\backslash b$	Word boundary
$\backslash B$	Non word boundary
$.$	Any character (sometimes excludes the end of line)
$^$	Start of input (or start of line)
$\$$	End of input (or end of line)