LeanDog
AN AGILE SOFTWARE STUDIO

Regular Expression Quick Reference

Characters

Character	Description	Example
Any character except [\^$.|?*+()	All characters except the listed special characters match a single instance of themselves. { and } are literal characters, unless they're part of a valid regular expression token (e.g. the {n} qualifier).	a matches a
\ (backslash) followed by any of [\^$.|?*+(){}	A backslash escapes special characters to suppress their special meaning.	\+ matches +
\Q...\E	Matches the characters between \Q and \E literally, suppressing the meaning of special characters.	\Q+-*/\E matches +-*/
\n, \r, and \t	Matches a LF character, CR character and a tab character respectively, Can be used in character classes.	\r\n matches a DOS/Windows CRLF line break.

Character Classes or Character Sets

Character	Description	Example
[(open square bracket)	Starts a character class. A character class matches a single character out of all the possibilities offered by the character class. Inside a character class, different rules apply. The rules in this section are only valid inside character classes. The rules outside this sections are not valid in character classes, except for a few character escapes that are indicated with "can be used inside character classes".	
Any character except ^-]\ add that character to the possible matches for the character class	All characters except the listed special characters.	[abc] matches a, b, or c
\ (backslash) followed by any of ^-] \	A backslash escapes special characters to suppress their special meaning.	[\^\]] matches ^ or]
- (hyphen) except immediately after the opening [	Specifies a range of characters (Specifies a hyphen if placed immediately after the opening [)	[a-zA-Z0-9] matches any letter or digit
^ (caret) immediately after the opening [	Negates the character class, causing it to match a single character *not* listed in the character class. (Specifies a caret if placed anywhere except after the opening [)	[^a-d] matches x (any character except a, b, c, or d)
\d, \w and \s	Shorthand character classes matching digits, word characters (letters, digits, and underscores), and whitespace (spaces, tabs, and line breaks). Can be used inside and outside character classes.	[\d\s] matches a character that is a digit or whitespace

Dot

Character	Description	Example
. (dot)	Matches any single character except line break characters \r and \n.	. matches x or (almost) any other character

Anchors

Character	Description	Example
^ (caret)	Matches at the start of the string the regex pattern is applied to. Matches a position rather than a character.	^. matches a in abcdef
$ (dollar)	Matches at the end of the string the regex pattern is applied to. Matches a position rather than a character. Also matches before the very last line break if the string ends with a line break.	.$ matches f in abcdef

Word Boundaries

Character	Description	Example
\b	Matches at the position between word characters (anything matched by \w) and a non-word character (anything matched by [^\w] or \W as well as at the start and/or end of the string if the first and/or last characters in the string are word characters.	.\b matches c in abc
\B	Matches at the position between two word characters (i.e. the position between \w\w as well as the position between two non-word characters (i.e. \W\W).	\B.\B matches b in abc

Alternation

Character	Description	Example
l (pipe)	Causes the regex to match either the part on the left side, or the part on the right side. Can be strung together into a series of options	abcldeflxyz matches abc, def, or xyz
l (pipe)	The pipe has the lowest precedence of all operators. Use grouping to alternate only part of the regular expression.	abc(deflxyz) matches abcdef or abcxyz

Quantifiers

Character	Description	Example
? (question mark)	Makes the preceding item optional.	abc? matches ab or abc
* (star)	Repeats the previous item zero or more times.	".*" matches "def" "ghi" in abc "def" "ghi" jkl
+ (plus)	Repeats the previous item once or more times.	".+" matches "def" "ghi" in abc "def" "ghi" jkl
{n} where n is an integer >= 1	Repeats the previous item exactly n times.	a{3} matches aaa
{n,m} where n >= 0 and m >= n	Repeats the previous item between n and m times.	a{2,4} matches aaaa, aaa, or aa
{n,} where n >= 0	Repeats the previous item at least n times	a{2,} matches aaaaa in aaaaa