Sublime Text Regular Expression Cheat Sheet

2019-02-28·technique·614 words ·3 mins read ·94638 times read

CONTENTS

A cheat sheet about regular expressions in Sublime Text.

Special characters

expression	Description
`.`	Match any character
`^`	Match line begin
`$`	Match line end
`*`	Match previous RE 0 or more times greedily
`*?`	Match previous RE 0 or more times non-greedily
`+`	Match previous RE 1 or more times greedily
`+?`	Match previous RE 1 or more times non-greedily
`?`	Match previous RE 0 or 1 time greedily
`??`	Match previous RE 0 or 1 time non-greedily
`A|B`	Match either RE A or B
`{m}`	Match previous RE exactly m times
`{m,n}`	Match previous RE m to n times greedily
`{m, n}?`	Match previous RE m to n times, no-greedily

Character set

expression	Description
`[abc]`	Match either `a`, `b` or `c`
`[^abc]`	Match any character not in this set (i.e., not `a`, `b` and `c`)
`[a-z]`	Match the range from `a` to `z`
`[a-f2-8]`	Match the range from `a` to `z` or the range from `2` to `8`
`[a\-z]`	Match `a`, `-` or `z`
`[a-]`	Match `a`, `-`
`[-a]`	Match `-`, `a`
`[-a]`	Match `-`, `a`
`[{}*|()[]+\^$.?]`	Match either one of the chacters in `[]{}*|()+^$?.`

- Note that you can also use character class inside `[]`, for example, `[\w]` matches any character in `word` character class.

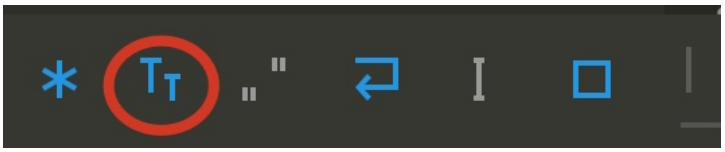
Character class

"Multiple character" character class

An expression of the form `[[:name:]]` matches the named character class `name`.

class name	Description
`alnum`	Any alpha-numeric character
`alpha`	Any alphabetic character.
`digit`	Any decimal digit.
`xdigit`	Any hexadecimal digit character.
`lower`	Any lower case character.
`upper`	Any upper case character.
`cntrl`	Any control character[1].
`print`	Any printable character.
`punct`	Any punctuation character. [2]
`space`	Any whitespace character. [3]
`word`	Any word character (alphanumeric characters plus the underscore).

Note: To use `upper` and `lower`, you have to enable case sensitve search.



"Single character" character class

class name	Description
`\d`	Equal to `[[:digit:]]`
`\l`	Equal to `[[:lower:]]`
`\u`	Equal to `[[:upper:]]`
`\s`	Equal to `[[:space:]]`
`\w`	Equal to `[[:word:]]`
`\D`	Equal to `[^[:digit:]]`
`\L`	Equal to `[^[:lower:]]`
`\U`	Equal to `[^[:upper:]]`
`\W`	Equal to `[^[:word:]]`

Regex groups

Defining capture groups

expression	Description
`(?<NAME>pattern)`	Define a regex group named `NAME` which you can later refer to with `\g{NAME}`
`(?=pattern)`	Positive lookahead, consumes zero characters, the preceding RE only matches if this matches
`(?!pattern)`	Negative lookahead, consumes zero characters, the preceding RE only matches if this does not match
`(?<=pattern)`	Positive lookbehind, consumes zero characters, the following RE will only match if preceded with this fixed length RE.
`(?<!pattern)`	Negative lookbehind, consumes zero characters, the following RE will only match if not preceded with this fixed length RE.

Refering to matching groups (capture groups)

expression	Description
`\1`	Refer to first regex group
`\g{1}`	Refer to first regex group

expression	Description
`\g{12}`	Refer to 12th regex group
`\g{-1}`	Refer to last regex group
`\g{-2}`	Refer to last but one regex group

- The regex groups are indexed by the order of their opening braces.
- Note the `\g{NUM}` form allows for matching regex group index larger than 9, for example, `\g{12}`.

Miscellaneous

Escapes

class name	Description
`\xdd`	A hexadecimal escape sequence - matches the single character whose code point is `0xdd`.
`\x{dddd}`	A hexadecimal escape sequence - matches the single character whose code point is `0xdddd`.

Word boundaries

The following escape sequences match the boundaries of words:

class name	Description
`\<`	Matches the start of a word.
`\>`	Matches the end of a word.
`\b`	Matches a word boundary (the start or end of a word).
`\B`	Matches only when not at a word boundary.

References

- https://www.boost.org/doc/libs/1_44_0/libs/regex/doc/html/boost_regex/syntax/perl_syntax.html
- http://docs.sublimetext.info/en/latest/search_and_replace/search_and_replace_overview.html
- https://github.com/stajahlee/sublime-cheat-sheets/blob/master/cheat-sheets/Regular%20Expressions.cheatsheet
- Character class

The title image is taken from here.

---

1. Control character explanation: https://en.wikipedia.org/wiki/Control_character ↩
2. There are 14 punctuation marks in English: https://grammar.yourdictionary.com/punctuation/what/fourteen-punctuation-marks.html ↩
3. For whitespace character, see https://en.wikipedia.org/wiki/Whitespace_character ↩

Author : jdhao
LastMod : 2020-12-23

Reward

#Sublime-Text #Regex

Manipulating Images with Alpha Channels in Pillow
Cropping Rotated Rectangles from Image with OpenCV