

Regular Expressions Cheat Sheet

Wildcards	
.	Any character
\w	Word character: Letters, numbers, and _
\W	Non-word character
\s	White space: Space, tab, or end of line (EOL)
\S	Non-white space character
\d	Digit character (but not . in 2.4)
\D	Non-digit character
\t	Tab character
\n	Newline character (some use \r)
\\	\ is the escape character so \\ matches a \, \^ matches ^, etc.

Repetitions	
*	Match 0 or more times
+	Match 1 or more times
?	Non-greedy-get shortest match
{n}	Match exactly n times
{n,}	Match at least n times
{n,m}	Match at least n, but not more than m times

Character classes	
[AEIOU]	Match 1 vowel character
[A-Z]	Match on capital letter
[A-Za-z0-9_]	Same as \w
^[A-Za-z]	Match one letter at beginning of string/line
[^A-Za-z]	Match any character that is not a letter (including EOL)—NOTE the very different meaning of ^ here. Within [], ^ means NOT.

Boundary qualifiers	
^	Match at start of string/line
\$	Match at end of string/line
^ABC\$	Exact match to ABC—Both starts and ends in ABC

Capturing and replacing	
()	Capture the search results between the parentheses and save for use. Save in order of appearance of (if multiple or nested.
\$1	Variables containing the captured text numbered in order of (. Can be used for replacing. Some editors (BBEdit) use \1, \2, etc.
\$2	

Some Examples	
<i>This is 1.0 (one) line to play with</i>	
Pattern	Matches:
\w*	<i>This</i>
\d+	<i>1</i>
[0-9\.\.]+	<i>1.0</i>
((\w*)\s(\w*))	\$1= <i>This is</i> , \$2= <i>This</i> , \$3= <i>is</i> (or \1,\2,\3 in BBEEdit)
\w*\$	<i>With</i>
"is "{2}	<i>is is</i> (from This is)
\(.+\)	<i>(one)</i> (note the escaping of parentheses)
\w+?	<i>T</i> (note the non-greediness of ?)
\w*?	Blank match (why?)