Table of Contents

Appendix E. Regular Expressions.

Appendix E. Regular Expressions.

Summary of regular-expression constructs

Construct	Matches	
	Characters	
X	The character x	
//	The backslash character	
\0n	The character with octal value $0n (0 \le n \le 7)$	
\0nn	The character with octal value $0nn (0 \le n \le 7)$	
\0mnn	The character with octal value 0mnn (0 \leq m \leq 3, 0 \leq n \leq 7)	
\xhh	The character with hexadecimal value 0xhh	
\uhhhh	The character with hexadecimal value 0xhhhh	
\t	The tab character ('\u0009')	
\n	The newline (line feed) character ('\u000A')	
\r	The carriage-return character ('\u000D')	
\f	The form-feed character ('\u000C')	
\a	The alert (bell) character ('\u0007')	
\e	The escape character ('\u001B')	
\cx	The control character corresponding to x	
Character classes		
[abc]	a, b, or c (simple class)	
[abc]	Any character except a, b, or c (negation)	
[a-zA-Z]	a through z or A through Z, inclusive (range)	
[a-d[m-p]]	a through d, or m through p: [a-dm-p] (union)	
[a-z&&[def]]	d, e, or f (intersection)	
[a-z&&[^bc]]	a through z, except for b and c: [ad-z] (subtraction)	
[a-z&&[^m-p]]	a through z, and not m through p: [a-lq-z](subtraction)	
	Predefined character classes	
	Any character (may or may not match line terminators)	
\d	A digit: [0-9]	
\D	A non-digit: [^0-9]	
\s	A whitespace character: [\t\n\x0B\f\r]	
\S	A non-whitespace character: [^\s]	
\w	A word character: [a-zA-Z_0-9]	
\W	A non-word character: [^\w]	
	POSIX character classes (US-ASCII only)	
\p{Lower}	A lower-case alphabetic character: [a-z]	
\p{Upper}	An upper-case alphabetic character:[A-Z]	
\p{ASCII}	All ASCII:[\x00-\x7F]	
\p{Alpha}	An alphabetic character:[\p{Lower}\p{Upper}]	
\p{Digit}	A decimal digit: [0-9]	
\p{Alnum}	An alphanumeric character:[\p{Alpha}\p{Digit}]	
\p{Punct}	Punctuation: One of !"#\$%&'()*+,/:;<=>?@[\]^_`{ }~	
\p{Graph}	A visible character: [\p{Alnum}\p{Punct}]	

Construct	Matches	
	Characters	
\p{Print}	A printable character: [\p{Graph}]	
\p{Blank}	A space or a tab: [\t]	
\p{Cntrl}	A control character: [\x00-\x1F\x7F]	
\p{XDigit}	A hexadecimal digit: [0-9a-fA-F]	
\p{Space}	A whitespace character: [\t\n\x0B\f\r]	
Classes for Unicode blocks and categories		
\p{InGreek}	A character in the Greek block (simple block)	
\p{Lu}	An uppercase letter (simple category)	
\p{Sc}	A currency symbol	
\P{InGreek}	Any character except one in the Greek block (negation)	
[\p{L}&&[^\p{Lu}]]	Any letter except an uppercase letter (subtraction)	
Boundary matchers		
^	The beginning of a line	
\$	The end of a line	
\b	A word boundary	
\B	A non-word boundary	
\A	The beginning of the input	
\G	The end of the previous match	
\Z	The end of the input but for the final terminator, if any	
\z	The end of the input	
	Greedy quantifiers	
X?	X, once or not at all	
X*	X, zero or more times	
X+	X, one or more times	
X{n}	X, exactly n times	
X{n,}	X, at least n times	
X{n,m}	X, at least n but not more than m times	
	Reluctant quantifiers	
X??	X, once or not at all	
X*?	X, zero or more times	
X+?	X, one or more times	
X{n}?	X, exactly n times	
X{n,}?	X, at least n times	
X{n,m}?	X, at least n but not more than m times	
	Possessive quantifiers	
X?+	X, once or not at all	
X*+	X, zero or more times	
X++	X, one or more times	
X{n}+	X, exactly n times	
X{n,}+	X, at least n times	
X{n,m}+	X, at least n but not more than m times	
Logical operators		
XY	X followed by Y	
X Y	Either X or Y	
	I · · · ·	

Construct	Matches	
Characters		
(X)	X, as a capturing group	
Back references		
\n	Whatever the n th capturing group matched	
Quotation		
\	Nothing, but quotes the following character	
\Q	Nothing, but quotes all characters until \E	
\E	Nothing, but ends quoting started by \Q	
Special constructs (non-capturing)		
(?:X)	X, as a non-capturing group	
(?idmsux-idmsux)	Nothing, but turns match flags on - off	
(?idmsux-idmsux:X)	X, as a non-capturing group with the given flags on - off	
(?=X)	X, via zero-width positive lookahead	
(?!X)	X, via zero-width negative lookahead	
(?<=X)	X, via zero-width positive lookbehind	
(? X)</td <td>X, via zero-width negative lookbehind</td>	X, via zero-width negative lookbehind	
(?>X)	X, as an independent, non-capturing group	

From:

http://www.awareim.com/dokuwiki/ - **Documentation**

Permanent link:

http://www.awareim.com/dokuwiki/docs/8000_appendices/0500_regular_expressions?rev=1661083826

Last update: 2022/09/13 18:11

