

Regular expressions (RE)

basic BRE
extended ERE

Atoms

(re) matches a match for re (re is any regular expression) with the match noted for possible reporting

[chars]

a bracket expression, matching any one of the chars

.

matches any single character

\k matches the non-alphanumeric character k taken as an ordinary character, e.g. \\ matches a backslash character

\c where c is alphanumeric (possibly followed by other characters), an escape (AREs only),

{ when followed by a character other than a digit, matches the left-brace character “{”; when followed by a digit, it is the beginning of a bound

x where x is a single character with no other significance, matches that character.

Bracket

[0123456789] any of the characters

[0-9a-zA-Z] any characters in the range

[^0-9a-zA-Z] negation

Character classes

[:name:]

alpha A letter.

upper An upper-case letter.

lower A lower-case letter.

digit A decimal digit.

`xdigit` A hexadecimal digit.

`alnum` An alphanumeric (letter or digit).

`print` A "printable" (same as `graph`, except also including space).

`blank` A space or tab character.

`space` A character producing white space in displayed text.

`punct` A punctuation character.

`graph` A character with a visible representation (includes both `alnum` and `punct`).

`cntrl` A control character.

Escape characters

`\a` alert (bell) character, as in C

`\b` backspace, as in C

`\f` formfeed, as in C

`\n` newline, as in C

`\r` carriage return, as in C

`\t` horizontal tab, as in C

`\v` vertical tab, as in C are all available.

`\0` the character whose value is 0

Class-escapes

`\d` `[[:digit:]]`

`\s` `[[:space:]]`

`\w` `[[:alnum:]]_`

`\D` `[^[:digit:]]`

`\S` `[^[:space:]]`

`\W` `[^[:alnum:]]_`

Quantifiers

- * a sequence of 0 or more matches of the atom
- + a sequence of 1 or more matches of the atom
- ? a sequence of 0 or 1 matches of the atom
- {m} a sequence of exactly m matches of the atom
- {m,} a sequence of m or more matches of the atom
- {m,n} a sequence of m through n (inclusive) matches of the atom; m may not exceed n
- *? +? ?? {m}? {m,}? {m,n}? non-greedy

Anchors

- ^ matches at the beginning of a line
- \$ matches at the end of a line