

Extended Regular Expressions

Elementary Operations

single characters

except “ \ . \$ ^ [] - ? * + | () / { } % < > ”

concatenation (put characters together)

alternation (a|b|c)

[ab] == a|b

[a-k] == a|b|c|...|i|j|k

[a-z0-9] == any letter or digit

[^a] == any character but a

. matches any character except the newline

* -- Kleene Closure

+ -- Positive Closure

Examples:

[0-9]+ "[0-9]+

note: without the quotes it could be any character

[\t]+ -- is whitespace

(except CR).

There is a blank space character before the \t

Special Characters:

. -- matches any single character (except newline)

“ and \ -- quote the part as text

\t -- tab

\n -- newline

\b -- backspace

\ " -- double quote

\\ -- \

? -- this means the preceding was optional

ab? == a|ab

(ab)? == ab|e

^ -- means at the beginning of the line (unless it is inside of a [])

\$ means at the end of the line, same as \n

[^] -- means anything except

\ "[^"]*" is a double quoted string

{n,m} – n through m occurrences

a{1,3} is a or aa or aaa

{varName} – translation of varName from definition

/ -- matches only if followed by right part of /

0/1 means the 0 of 01 but not 02 or 03 or ...

() – grouping