



Regular Expression

EE 599

*Department of Electrical Engineering
University of Southern California*

Regular Expressions

In computing, a regular expression, also referred to as "regex" or "regexp", provides a concise and flexible means for matching strings of text, such as particular characters, words, or patterns of characters. A regular expression is written in a formal language that can be interpreted by a regular expression processor.

http://en.wikipedia.org/wiki/Regular_expression

Regular Expressions

Really clever "wild card" expressions for matching and parsing strings.

http://en.wikipedia.org/wiki/Regular_expression



Really smart “Find” or “Search”

Understanding Regular Expressions

- ❑ Very powerful and quite cryptic
- ❑ Fun once you understand them
- ❑ Regular expressions are a language unto themselves
- ❑ A language of "marker characters" - programming with characters
- ❑ It is kind of an "old school" language - compact

WHENEVER I LEARN A NEW SKILL I CONCOCT ELABORATE FANTASY SCENARIOS WHERE IT LETS ME SAVE THE DAY.

OH NO! THE KILLER MUST HAVE FOLLOWED HER ON VACATION!



BUT TO FIND THEM WE'D HAVE TO SEARCH THROUGH 200 MB OF EMAILS LOOKING FOR SOMETHING FORMATTED LIKE AN ADDRESS!



IT'S HOPELESS!

EVERYBODY STAND BACK.



I KNOW REGULAR EXPRESSIONS.



<http://xkcd.com/208/>

Regular Expressions – Quick Guide (1)

- .** - Any Character Except New Line
- \d** - Digit (0-9)
- \D** - Not a Digit (0-9)
- \w** - Word Character (a-z, A-Z, 0-9, _)
- \W** - Not a Word Character
- \s** - Whitespace (space, tab, newline)
- \S** - Not Whitespace (space, tab, newline)

Regular Expressions – Quick Guide (2)

- \b** - Word Boundary
- \B** - Not a Word Boundary
- ^** - Beginning of a String
- \$** - End of a String

Regular Expressions – Quick Guide (3)

- [] - Matches Characters in brackets
- [^] - Matches Characters NOT in brackets
- | - Either Or
- () - Group

Regular Expressions – Quick Guide (4)

Quantifiers:

- * - 0 or More
- + - 1 or More
- ? - 0 or One
- {3} - Exact Number
- {3,4} - Range of Numbers (Minimum, Maximum)

Regular Expressions – Quick Guide (5)

Sample Reg-Ex

`[a-zA-Z0-9_+.]+@[a-zA-Z0-9-]+\.[a-zA-Z0-9-.]+`

Thank You!