

Regular Expressions

CS 272 Software Development

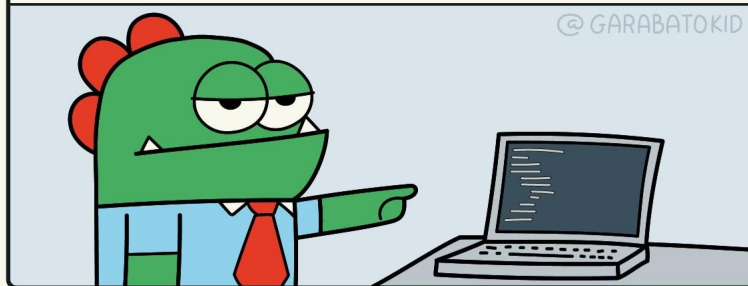
Regular Expressions

- Often abbreviated as **regex**
- Standard* way to express and match String patterns
 - *e.g. what an email looks like, find all emails on web*
- Supported by most programming languages
- Originally a compact representation of regular languages



HOW TO REGEX

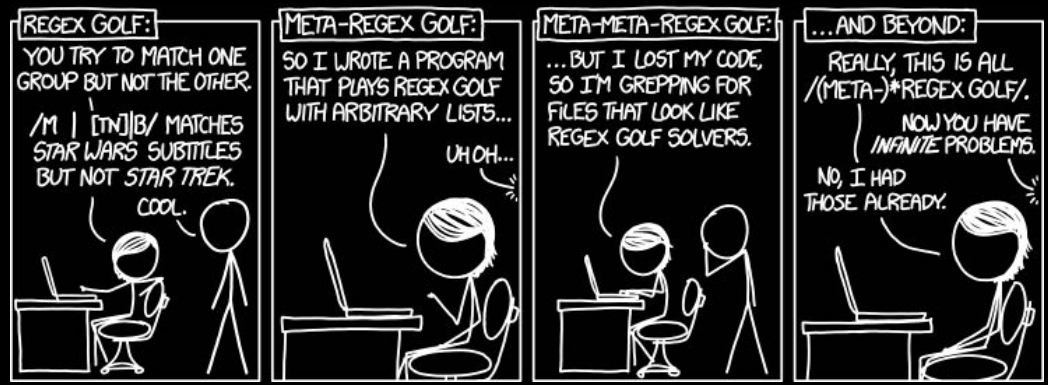
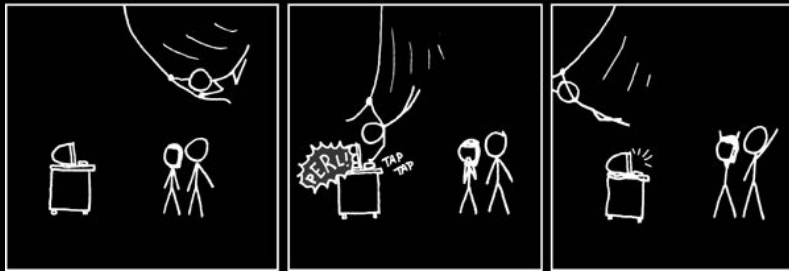
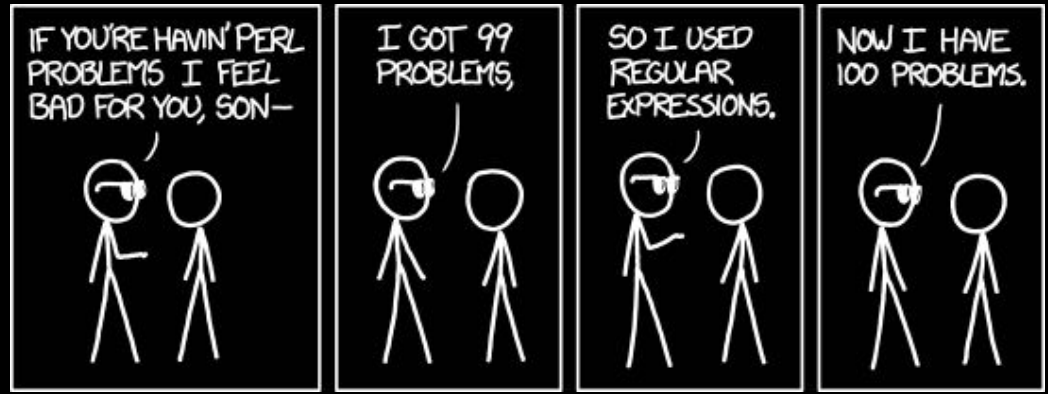
STEP 1: OPEN YOUR FAVORITE EDITOR



STEP 2: LET YOUR CAT PLAY ON YOUR KEYBOARD



<https://twitter.com/garabatokid/status/1147063121678389253>



<https://xkcd.com/208> · <https://xkcd.com/1171> · <https://xkcd.com/1313/>

`< / ? \ w + \ b [^ >] * >`

- `<` match `<` character
- `/?` match `/` 0 or 1 times
- `\w+` match 1 or more word characters (letters, `_`)
- `\b` match a boundary
- `[^>]*` match any char except `>` 0 or more times
- `>` match the `>` character

<https://regex101.com/r/DDQcFj/1>



`< / ? \ w + \ b [^ >] * >`

- `<` match `<` character
- `/?` match `/` 0 or 1 times
- `\w+` match 1 or more word characters (letters, `_`)
- `\b` match a boundary
- `[^>]*` match any char except `>` 0 or more times
- `>` match the `>` character

<https://regex101.com/r/DDQcFj/1>



Regular Expressions in Java

- Java Strings use `\` backslash for special characters
 - e.g. `\n` for newline, `\t` for tab, etc.
- Regexes also use `\` backslash for special characters
 - e.g. `\s` for whitespaces

```
String hello = "hello\tworld"; // hello    world
```

```
String regex = "hello\\sworld"; // hello\sworld
```

<https://docs.oracle.com/javase/specs/jls/se17/html/jls-3.html#jls-EscapeSequence>



Common Regexes

- **\b** word boundary
- **\s** whitespace character
- **\t** tab
- **\n** newline
- **\r** carriage return
- ...

<https://docs.oracle.com/en/java/javase/17/docs/api/java.base/java/util/regex/Pattern.html>

Escape Sequences

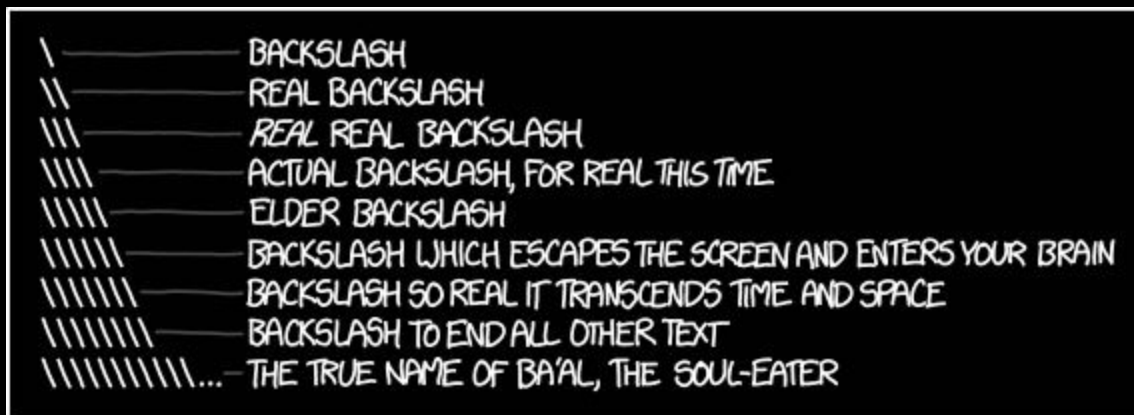
- **\b** backspace
- **\s** space
- **\t** tab
- **\n** newline
- **\r** carriage return
- ...



<https://docs.oracle.com/javase/specs/jls/se17/html/jls-3.html#jls-3.10.7>



Regular Expressions in Java



<https://xkcd.com/1638>



Regular Expressions in Java

- Express regular expression as String literal
- Use [Pattern](#) to compile into regex
- Use [Matcher](#) to match against that regex
- Various other convenience methods as well
 - e.g. [String.replaceAll\(regex, replacement\)](#)

<https://www.cs.usfca.edu/~cs272/javadoc/api/java.base/java/util/regex/Pattern.html>



Java Regex Resources

The Java Tutorials - Lesson: Regular Expressions

<https://docs.oracle.com/javase/tutorial/essential/regex/index.html>

Pattern Class

<https://docs.oracle.com/en/java/javase/17/docs/api/java.base/java/util/regex/Pattern.html>

Regex101 Editor - Flavor: Java 8

<https://regex101.com/>



Questions?

