Test Driven Regex

Stan Kurdziel
Mobile Developer at ResMed
@akasek
What is regex?

Regular Expressions - Notation to express string matching

Starts simple:

<table>
<thead>
<tr>
<th>notation</th>
<th>meaning</th>
<th>example</th>
<th>matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>1 or none</td>
<td>colou?r</td>
<td>color, colour</td>
</tr>
<tr>
<td>+</td>
<td>1 or many</td>
<td>ex+tra</td>
<td>extra, exxtra, exxxtra</td>
</tr>
<tr>
<td>*</td>
<td>0 or many</td>
<td>o*ps</td>
<td>ps, ops, oops, ooooooops</td>
</tr>
<tr>
<td>{n}</td>
<td>repeats n times</td>
<td>ba(na){2}</td>
<td>banana</td>
</tr>
</tbody>
</table>
The problem with regex

Some people, when confronted with a problem, think “I know, I'll use regular expressions.” Now they have two problems.

--Jamie Zawinski (an early Netscape engineer)

Incidentally he copied earlier quotes with ‘sed’ and ‘awk’ instead of ‘regex’
They tend to get *complicated* very quickly!

And how about readability? What does this mean?

```regex
(?<=\.) \{2,}(?=\[A-Z])
```

https://docs.oracle.com/javase/10/docs/api/java/util/regex/Pattern.html
There are some tools available that help

Ex: https://www.freeformatter.com/java-regex-tester.html

Java Regular Expression:

(?<=\.) \{2,\}?=[A-Z]

Entry to test against:

Let's be clear. The planet is not in jeopardy. We are in jeopardy. We haven't got the power to destroy the planet - or to save it. But we might have the power to save ourselves.

Many others too... Having tools is nice, but in my opinion, these don't bring the readability / maintainability up to par with "normal" code
How can we write readable / maintainable regex?

Phone number parsing: 858-555-5555


Off to the code… http://github.com/sek/tdd-regex