SY306 Web and Databases for Cyber Operations
SlideSet #7: Regular expressions
(JavaScript and Python)

http://www.w3schools.com/jsref/jsref_obj_regexp.asp

Regular Expressions

- Describe a pattern of characters
- JS: /pattern/modifiers

Is this a proper email address?
email = “instructor@usna.edu”

Regular Expressions are powerful
if( email.search(/\w+@\w+\.\w\w/) > -1){
    alert “Valid email!”;
}


Some Regular Expression Quantifiers and Metacharacters

<table>
<thead>
<tr>
<th>Quantifier/Symbol</th>
<th>Matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>{X}</td>
<td></td>
</tr>
<tr>
<td>{X,Y}</td>
<td></td>
</tr>
<tr>
<td>{X,}</td>
<td></td>
</tr>
<tr>
<td>+</td>
<td></td>
</tr>
<tr>
<td>*</td>
<td></td>
</tr>
<tr>
<td>?</td>
<td></td>
</tr>
<tr>
<td>^</td>
<td></td>
</tr>
<tr>
<td>$</td>
<td></td>
</tr>
<tr>
<td>\b</td>
<td></td>
</tr>
<tr>
<td>\w</td>
<td></td>
</tr>
<tr>
<td>\d</td>
<td></td>
</tr>
<tr>
<td>\s</td>
<td></td>
</tr>
<tr>
<td>\S</td>
<td></td>
</tr>
<tr>
<td>[abc]</td>
<td></td>
</tr>
</tbody>
</table>

Modifiers

- /pattern/modifiers

- Modifiers
  - i
  - g
  - m
  - x (Python)
Regular Expressions – JavaScript Examples

```javascript
<script>
var myString = "Now is is the time";
alert ("Test string is " + myString + ":");

if (myString.search(/Now/) > -1) alert ('Search 1 success');

searchPos = myString.search(/^Now/);
if (searchPos > -1) alert ('Search 2 success');

searchRes = /Now$/.test(myString);
if (searchRes) alert ('Search 3 success');

pattern = /\b(w+OW)\b/i;
searchRes = pattern.test(myString);
if (searchRes) alert ('Search 4 success');

pattern = /\b(w+)\s\d\b/;
searchMatch = pattern.exec(myString)
if (searchMatch)
    alert ('Search 5 success and returned array of size ' + searchMatch.length + ' with content ' + searchMatch)
</script> ...
```

Regular Expressions – Python Examples

```python
#!/usr/bin/env python3
import re
myString = "Now is is the time";
print ("Test string is " + myString + "")
if re.search(r'Now', myString) :
    print ('Search 1 success')
searchObj = re.search(r'^Now', myString)
if searchObj:
    print ('Search 2 success')
searchObj = re.search(r'Now$', myString)
if searchObj:
    print ('Search 3 success')
searchObj = re.search(r'\b ( \w+ OW ) \b', myString, re.X)
if searchObj:
    print ('Search 4 success: ' + searchObj.group(1))
searchObj = re.search(r'\b ( \w+ ) \s ( \d )\b', myString, re.X)
if searchObj:
    print ('Search 5 success: ' + searchObj.group(1) + " + " + searchObj.group(2))
```
Search and Replace

• JS: newstring = oldstring.replace(pattern, replacement)
• Python: newstring = re.sub(pattern, replacement, oldstring, max=0)

• Example (replace aa with bb):
  • line = “This string has aa here and aa here”
  • JS: newline = line.replace(/aa/g,”bb”)
  • Python: newline = re.sub(r’a’a’, ’bb’, line)

Exercise #1

• Write the expression to replace one or more newline characters in a string with “&&”.
• Make it work for both Unix (\n) and Windows (\r\n)
Uses for Regular Expressions

- Input validation
- Input sanitization