Adminstrivia

• Homework 6 out today
  – due Saturday night (by midnight)
Last Time

String: It's abnormal that cats and dogs are raining

Regex: \b\w+s\b

Global match (g):  

Click

dogs
Homework 6

• From UIUC POS Tagger demo: sample.txt

Helicopters will patrol the temporary no-fly zone around New Jersey's MetLife Stadium Sunday, with F-16s based in Atlantic City ready to be scrambled if an unauthorized aircraft does enter the restricted airspace.

Down below, bomb-sniffing dogs will patrol the trains and buses that are expected to take approximately 30,000 of the 80,000-plus spectators to Sunday's Super Bowl between the Denver Broncos and Seattle Seahawks.

The Transportation Security Administration said it has added about two dozen dogs to monitor passengers coming in and out of the airport around the Super Bowl.

On Saturday, TSA agents demonstrated how the dogs can sniff out many different types of explosives. Once they do, they're trained to sit rather than attack, so as not to raise suspicion or create a panic.

TSA spokeswoman Lisa Farbstein said the dogs undergo 12 weeks of training, which costs about $200,000, factoring in food, vehicles and salaries for trainers.

Dogs have been used in cargo areas for some time, but have just been introduced recently in passenger areas at Newark and JFK airports. JFK has one dog and Newark has a handful, Farbstein said.
Homework 6

For each question, provide the screen snapshot with the regex and results.
Homework 6

• Question 1: write a regex that finds all the acronyms in the article.
Homework 6

• Question 2: write a regex that finds all the **numeric items** in the article.
Homework 6

- Question 3: write a regex that finds all **Noun-Noun compounds**

![Diagram of regex examples](Image)
Homework 6

• Question 4: write a regex that finds all the main verbs (exclude auxiliaries) in the article.

Note: search may return an array with submatches: ok if main verb is a submatch,
e.g. will patrol, will, patrol

[match]  [submatches]
Homework 6

• Question 5: write a regex that finds all the passive verbs

String: Helicopters will patrol the temporary no-fly zone arou
Regex: ____________________________ Global match (g): ✓  Click

be scrambled, be, scrambled
are expected, are, expected
been used, been, used
been introduced, been, introduced

answer as a submatch ok
Javascript Regexp Tester with Replace

http://dingo.sbs.arizona.edu/~sandiway/ling508-15/rep-test.html
Suppose we want to modify string str => modified_str

We'll need the string method replace():

- var regex = new RegExp(re_s,flag_s);
- var modified_str = str.replace(regex,replacement)
- replacement string can contain $n
- (n = group number)

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Inserts</th>
</tr>
</thead>
<tbody>
<tr>
<td>$$</td>
<td>Inserts a &quot;$&quot;.</td>
</tr>
<tr>
<td>$$</td>
<td>Inserts the matched substring.</td>
</tr>
<tr>
<td>$`</td>
<td>Inserts the portion of the string that precedes the matched substring.</td>
</tr>
<tr>
<td>$'</td>
<td>Inserts the portion of the string that follows the matched substring.</td>
</tr>
<tr>
<td>$n or $nn</td>
<td>Where n or nn are decimal digits, inserts the nth parenthesized submatch string, provided the first argument was a RegExp object.</td>
</tr>
</tbody>
</table>
Javascript Regexp Tester **with Replace**

```html
22 <body>
23 <h1> Javascript Regex Replace </h1>
24 <form>
25 String: <input type=text name=str size=30>
26 <br>
27 Regex: <input type=text name=re size=30>
28 <br>
29 Replace: <input type=text name=rp size=30>
30 <br>
31 Global match (g): <input type=checkbox name=g value=g>
32 <input type=button value="Click" onclick="f(this)">
33 </form>
34 <script>
35 document.getElementsByTagName("form")[0].re.value = "Mr\\. ([A-Z][a-z]*)"
36 </script>
37 <div id="output"></div>
38 </body> </html>
```
Javascript Regexp Tester with Replace

5. `<script>`
6. function f(e) {
7.     var o = document.getElementById("output");
8.     o.innerHTML = "";
9.     var re_s = e.form.re.value;
10.    var s = e.form.str.value;
11.    var r = e.form.rp.value;
12.    if (re_s !== "") {
13.        var flag_s = "";
14.        if (e.form.g.checked) {
15.            flag_s += "g"
16.        }
17.        var regex = new RegExp(re_s,flag_s);
18.        o.innerHTML = s.replace(regex,r).toString()
19.    }
20. }
21. `</script>`
Example with HTML replace

- **Example**: replacement containing HTML code
  - string to be modified: `[c_Q[q[q][who]]][c_Q[c_Q][Tpast[q[q][who]]][Tpast[v_unerg][Tpast[Tpast][v_unerg[q[q][who]][v_unerg[v_unerg][laugh]]]]]]]
  - regex: `(_(.+?)('[\[\]]))`
  - replacement string: `<sub>$1</sub>$2`

```html
<!DOCTYPE html>
<html>
<body>
<p>This text contains <sub>subscript</sub> text.</p>
<p>This text contains <sup>superscript</sup> text.</p>
</body>
</html>
```

This text contains `subscript` text.
This text contains `superscript` text.

http://www.w3schools.com
Javascript Regexp Tester with Replace

Javascript Regex Replace

String: [c_Q[q[q][who]]c_Q[c_Q][Tpast[q[q][who]][Tpast[v_unerg]
[Tpast[Tpast][v_unerg[q[q][who]][v_unerg[v_unerg][laugh]]]]]]

Regex: \(._{.+?}([{\}\[}]\)\)
Replace: \(<\text{sub}>\$1</\text{sub}>\$2\)
Global match (g): Click

[c_Q[q[q][who]]c_Q[c_Q][Tpast[q[q][who]][Tpast[v_unerg][Tpast[Tpast][v_unerg[q[q][who]][v_unerg[v_unerg][laugh]]]]]]}