Activity 2-6

March 15, 2012

Task 1: String Matching & Regular Expressions

The program `redemo.py` is a demo that Python provides.

1. Enter the following lines into the top box. What do you think brackets do?

   ```
   n
   fn
   [fn]
   [aeiou]
   ```

2. Get all the *words* that that append the letters `b`,`f`, or `m` to your name (or the suffix of your name).

3. Enter the following lines into the top box. What do you think `\w` does? What do you think the `+` sign does?

   ```
   f
   f\w
   f\w\w
   f\w+
   ```

4. Enter the following lines into the top box. What do you think `\s` does?

   ```
   \s
   \sm
   ```

5. Get all the words that start with `b`,`f`, or `m`.

6. Remember that \n is a special character that denotes a line break. Get all the words that appear at the *end* of a line.

7. Add some numbers to the middle box. Test what `\d` does.

Task 2: More Examples of Regular Expressions

Close the demo and return to `ACT2-7.py`. Download and save `poem.txt` in the same directory as `ACT2-6.py` (remember why we need to put `poem.txt` in the same directory as the python program?). Press `F5`.

1. Run the following statements and figure out what each line does.

```
1  myStr = readShel()
2
3  printRegEx('g\w+',myStr)
4  printRegEx('\sg\w+',myStr)
5  printRegEx('\s[gG]\w+',myStr)
6
7  printRegEx('out',myStr)
8  printRegEx('\sout',myStr)
9  printRegEx('\wout',myStr)
10 printRegEx('\w+out',myStr)
11 printRegEx('[\s\w]out',myStr)
12 printRegEx('[\s\w][Oo]ut',myStr)
```

2. Design regular expressions that print the following. Don't worry about capitalization (just get all instances of lowercase).

   (a) Print all occurrences of the substring `it`.

   (b) Print all occurrences of the **word** `it`. A word should be surrounded by whitespace. There are five instances of the word `it`.

   (c) Print all words that contain `it` and **at least one other letter**. There are six such words.

   (d) Print all words that end in `ing`. There are two such words.

   (e) Print all phrases surrounded by double quotes (all occurrences of speech). There is only one phrase.

   (f) Print all contractions (words with single quotes). Remember that single quotes are "special" - they require a \. There are two such words (`She'd` and `I'll`), but write the expression to return any contraction.