MIT AITI
Python Software Development

Lab 9: Regular Expressions

In this lab, you will learn how to search for content inside of a string using regular expressions.

1. This will be VERY useful in the lab, so at least skim over it. It's a thorough explanation of regular expressions in python. If you have any questions on this lab, PLEASE have this page open, as there's a good chance we will reference it:
   http://docs.python.org/library/re.html

2. Run reExamples.py and make sure you understand all the examples. Use the python documentation to help (link above).

3. Make a function isEmail(inp) that, when given a possible email string, returns True if the string matches the following: any amount of any case of letters, numbers, or underscores, followed by an @ symbol followed by any number of any case letters or numbers followed by at least one . followed by between 2 and 4 any case letters.
   ```
   isEmail('blah@hello.com') returns True
   isEmail('sd$sd@hello.com') returns False
   ```

4. Make a function getTxts(files) that, when given a list of space separated files, returns a list of filenames that have .txt at the end. The input list will always end in a space.
   ```
   getTxts('yo.html blah.txt woah.txt he ')
   returns ['blah.txt', 'woah.txt']
   ```

5. Make a function percAwesome(inp) that determines the percentage of awesomeness a string is by looking through the inp, which should be a space separated list of words, adding all the words that contain awesome or awes0me, and returning that percentage of these in the list of words, rounded to the nearest 10th of a percentage.
   isAwesome('iamawesomeblah and awes0me is as awesomeo does') returns 42.9 (3/7*100)
   isAwesome('hello my name is wayawesomedude') returns 20.0 (1/5*100)

6. CHALLENGE PROBLEM: Output the weather, in celcius from this website:
   http://www.wunderground.com/global/stations/65432.html
   Here are a few hints:

- Open up getWeather.py and run it; it should output the html from the url given.
- You only need to add ONE line to get this to work
- This line will have both the words "search" and "group"
- In order to determine what the regex should be, look in the html and find the place where the temperature is. Then look for identifying html that is unique and use trial and error to find something actually unique.
- In order to get a . character to match new lines (it doesn't by default), use the flag re.S
  - If this doesn't make sense, search for "re.DOTALL" inside this page: http://docs.python.org/library/re.html