LING 388: Computers and Language

Lecture 15
Today's Topics

• We didn't finish the exercises from last time ...
• But first the Windows IDLE
Python IDLE on Windows

- https://docs.python.org/3/faq/windows.html

First, you need to make sure that your command window recognises the word “python” as an instruction to start the interpreter. If you have opened a command window, you should try entering the command `python` and hitting return:

```
C:\Users\YourName> python

You should then see something like:
```
Python 3.3.0 (v3.3.0:bd8af90ebf2, Sep 29 2012, 10:55:48) [MSC v.1600 32 bit (In Type "help", "copyright", "credits" or "license" for more information.
>>>```

Add path to python executable to environment variable PATH:

_Control Panel / System / Advanced System Settings / Environmental Variables_
Python IDLE on Windows

[Images of System Properties, Environment Variables, and Edit environment variable windows with red arrows pointing to different sections.]

crashdeskggeek.com
Python IDLE on Windows

- `import os`
- `os.getcwd()`
- `os.chdir(r"..path..")`
  - Note: raw string (r) because of \ is directory separator in Windows and \ is the escape meta-character for strings.
Slides from last time ...

• Download `wordlist.py` (Brown Corpus words) to your computer and put in a directory where your Python can see it.

• **Best practice**: put it in the same directory as your Python working directory
More Python regex practice

• Run the following:

```python
>>> from wordlist import wordlist as wordlist
>>> len(wordlist)
56057
>>> 'python' in wordlist
True
>>> 'Python' in wordlist
False
>>> 'Monty' in wordlist
True
```
More Python regex practice

• Exercise 1: produce a list of all the words in wordlist that having two a's in a row

```python
aa = [word for word in wordlist if re.search('aa', word)]
len(aa)
```

• Exercise 2: are there more words with two b's in a row?

• Exercise 3: words with two p's or b's or d's in a row – which is the most frequent?
More Python regex practice

• Exercise 4: find a word with both bb and dd in it

• Exercise 5: are there any words with both pp and dd?

• Exercise 6: find words ending in zac. How many are there?
  • Recall: meta-character for the end of line anchor

• Exercise 7: find words beginning in anti. How many are there?
  • Hint: some cases may begin with a capital letter
Extra Credit: Python regex


• **Note**: The Brown Corpus is an American English corpus

• **Question 1**: Find four examples of words spelled both ways but present in the corpus

• **Question 2**: Consider the –or vs. –our ending, as in *color* and *colour*
  • Write Python code that searches wordlist and finds all examples present with both endings
  • Show your code and the result of your code running
Extra Credit: Python regex

• New deadline (by next Monday midnight) to TA Patricia Lee