

ME 499/599: Computer Programming for Mechanical Systems

Fall 2014

Lab 7: More War and Peace

Purpose

The goal of this lab is to give you some more experience working with some of the Python structures that we've looked at this term.

Preparation

Look over homework 2 again, and check out the reference solutions. We're going to use War and Peace again for this lab.

Assignment

1. Draw a histogram of the word lengths in War and Peace. Make sure the histogram reflects the lengths of words, not words and punctuation.
2. Draw a histogram of the letter frequencies in War and Peace.
3. Pick another large text from Project Gutenberg, and run the same analysis on that, using the same code.
4. Do it again, but this time for the French version of War and Peace (La Guerre et la Paix). Is French statistically different from English?
5. How many words do the English and French versions of War and Peace have in common? Count them and print them out.
6. Print out the n most common words in both English and French, alongside each other, in a format like this:

```
English French
=====
war guerre
```

Grading

If you use the built-in complex number implementation in your code, you will get zero points for this lab. The goal of this lab is for you to write your own code for complex numbers. You can, however, use the built-in implementation to test your own code.

There are 10 points available in this lab.

1. 1 point for correctly getting rid of punctuation in words. 1 point for using a sensible data structure to keep the counts. 1 point for getting the right counts. 1 point for a correctly-labeled histogram. [4 points total]
2. 1 point for correctly counting the letter frequencies. 1 point for a correctly-labeled histogram. [2 points total]
3. 1 point for the code working correctly on another book. [1 point total]
4. 1 point for the code working on the French version of War and Peace. You don't actually have to determine formally if French and English are different. [1 point total]
5. 1 point for counting the common words in the two versions. [1 point total]
6. 1 point for a well-formatted list of words, side-by-side. [1 point total]

Thoughts

1. Think about what data structure you should use to keep track of the word length and letter counts.
2. You should also think about how to encapsulate your code with classes and functions. Try to have each piece of code be as short and concise as possible. One potential design is to hide all of the implementation of the book in a class, so that you can do something like this:

```
war_and_peace = Book('wap.txt')
war_and_peace.DisplayWordHistogram()
war_and_peace.DisplayLetterFrequencies()

other_book = Book('other.txt')
other_book.DisplayWordHistogram()
```

The Rules

Everything you do for this lab should be your own work. Don't look up the answers on the web, or copy them from any other source. You can look up general information about Python on the web, but no copying code you find there. Read the code, close the browser, then write your own code.

Page written by [Bill Smart](#).