

# Worksheet #3: Inaugural Addresses

## CSC 143 – Computer Programming II

### Due Monday 1/30/2017

#### Information

To complete Worksheet #3, you will need to review the material in Chapter 10, which covers ArrayLists.

- Include an answer sheet in PDF format, in which you provide answers to the questions below. Submit worksheet 3 through Canvas. Submit any Java code that you wrote to answer the questions as a .java file.

#### Background

In this worksheet, you will be using lists to store text and analyze vocabulary. Specifically, you will be analyzing inaugural addresses of the Presidents of the United States. These files are text files on disk; your file will read them into memory, parse them into a list of words, and perform several operations.

- You will start by picking two inaugural addresses to compare.
- First, you will be implementing a function that reports the unique words in each of the two inaugural addresses you chose.
- Second, you will be implementing a function that computes the overlap between the two inaugural addresses – the number of words common to both speeches.
- Third, you will implement a program to read from a large number of text files and compile a cross-word matrix. This is a matrix where each row and column represents a different speech, and the value in the matrix represents the number of words shared between the two speeches (and 0 along the diagonal). For example, if we only look at the first 6 inaugural speeches, we have the following cross-word matrix:

```
["George Washington (1789)",  
"George Washington (1789)",  
"John Adams (1797)",  
"Thomas Jefferson (1801)",  
"Thomas Jefferson (1805)",  
"James Madison (1809)"]
```

```
[[ 0, 59, 244, 216, 219, 183]  
 [ 59, 0, 55, 55, 53, 46]  
 [244, 55, 0, 253, 279, 229]  
 [216, 55, 253, 0, 258, 217]  
 [219, 53, 279, 258, 0, 221]  
 [183, 46, 229, 217, 221, 0]]
```

For example: John Adams' 1797 inaugural address shared 244 words in common with George Washington's 1789 inaugural address; James Madison's 1809 inaugural address shared 217 words in common with Thomas Jefferson's 1801 inaugural address; etc.

## Assignment

Implement an object called `InauguralAnalysis` that contains all of the functionality required to answer the following questions. This object should implement a public static void `main` method, which you should utilize to answer the following questions, and use various fields.

1. Which two inaugural addresses did you choose to compare?  
Implement two `ArrayLists` as fields in your class. These should store the unique words in the two inaugural addresses that you chose.  
Implement a method to populate these two lists. When the lists are populated, print out the name of the President and the year of the inaugural address, as well as how many words.  
For example:

George Washington (1789): 654 unique words  
George Washington (1793): 92 unique words

2. Implement a method to compute the overlap in words between the two inaugural addresses that you selected. Find the number of unique words common to both speeches. For example:

George Washington (1789) and George Washington (1793) have 52 unique words in common.

3. Implement a method that will read through some or all of the inaugural addresses, and compile a cross-word matrix (described in the first page). This is a matrix where each row and column represents a different speech, and the value in the matrix represents the number of words shared between the two speeches (and 0 along the diagonal). Write this matrix to a CSV file.

Suggested approach: work out the algorithm for a smaller data set (the first 6 speeches) and compare to the numbers on the first page. Once it is ready, apply it to the entire data set.

## Hints

I will give you three hints to help you with this homework.

- Utilize the `metadata.csv` file to link each speech to the name of the President who gave it and the year, instead of copying-and-pasting data into your Java program.
- Utilize a Java CSV library to read and parse CSV files like `metadata.csv`.
- You are allowed to look ahead to Chapter 11, and implement these operations with `Set` objects.