

# **Programming Assignment #9**

## **CS253, Fall 2017**

### **Due: Tuesday, Dec. 5<sup>th</sup>**

## ***“The Efficiency Assignment”***

### ***Motivation & Description***

The task for programming assignment #9 is exactly the same as the task for PA8. The difference is in how it will be graded. Whereas in previous assignments 100% of the grade was determined by correctness, in PA9 50% of the grade is correctness and 50% is efficiency on correct cases. (Incorrect answers get no speed points.)

### ***Task***

Your program will take two filenames and two numbers as arguments, and write out a single filename to `std::cout`. The first input file is the document you are comparing to. It is a standard text document. The second file is more complicated. It contains a series of filenames separated by whitespace, so that they can be read using the `>>` operator. (We will probably put one file name per line, but your program shouldn't count on that.) The first filename in the second file is the name of the exceptions file. All other filenames in the second file are the names of background documents. The third and fourth arguments are numbers, specifically the lowest reading level (as defined by the Coleman-Liau Index) allowed and the highest reading level. It is an error if the third value is greater than the fourth.

Your program should analyze the reading level of every background document, as in PA7. Then, for documents whose reading level is within the target range (as determined by the 3<sup>rd</sup> and 4<sup>th</sup> arguments), compare them to the target document using TFIDF (as in PA8). Print to `std::cout` the filename of the most similar document to the target document that is within the target reading level range.

### ***Hints***

- Make sure your PA8 code is correct. Wrong answers get no points.
- Write a suite of test cases using GoogleTest. That way, when you work on your code to make it faster, you can double check that you aren't introducing new bugs.
- Use the valgrind profiler to find out where your program is spending its time. There is no point optimizing parts of your code that aren't significant time sinks.
- Often, the STL is the easiest way to make your code faster.

### ***Submitting Your Work***

To submit your program, first create a single tar file containing all of your source files (i.e. your `.cpp` and `.h` files) and your makefile, but not your compiled files (no executable or `.o` files, please). Then submit the tar file as PA9 on Canvas.

## ***Policies***

All work you submit must be your own. You may not submit code written by a peer, a former student, or anyone else. You may not copy or buy code from the web. The department academic integrity policies apply.

You may not submit your program late. To receive credit, it must be submitted on Tuesday, Nov. 9th. There is no late period. The exception is an unforeseeable emergency, for example a medical crisis or a death in the immediate family. If an unforeseeable emergency arises, talk to the instructor.