Data Structures and Algorithms, Term Project
Instructor : Assistant Prof. K. Egemen Ozden

-- Word Counter --

1. Main Requirements

You are expected to write a simple console application which reads a text file which can be piece of a novel, newsletter etc. Example data files will be in OIS. Text files will be simple like (much longer though):

--It was many and many a year ago,
In a kingdom by the sea,
That a maiden there lived whom you may know
By the name of Annabel lee;
And this maiden she lived with no other thought
Than to love and be loved by me.

I was a child and she was a child,
In this kingdom by the sea;
……..
--

Your program must be able to read individual words from such files and insert it into a ~~circular singly linked list~~, if it is a new word. If that word is already encountered before its counter must be increased by 1. The list must always be ordered alphabetically. In other words when inserting the new word, you must insert it to the right place so alphabetical order is not broken.

After reading and processing is over, the program gives two options to the user.

a. List the most popular words
b. List total number of words.
c. Print the counter value of a specific word.

If option a is selected, the program asks how many words does the user want. For example if user types 5, it will print the most popular 5 words in the text. If option b is selected the user will input the word she would like to be printed. For example, in the above sample text the word “by” occurs 4 times so this must be printed if the user asks the word “by”.

Whole application can be implemented with console facilities(ie you do not need advanced GUI elements). The project consists of two parts.

A. Implementation of a ~~circular singly linked list class~~ of “persons” with following operations : add to head, remove from head, add to tail, remove from tail, ~~insertOrderly~~ and getNodeByValue. InsertOrderly is used to enter a new element to the proper alphabetical place. getNodeByValue will be used to search the list word the query word, so its counter value can be printed.

This will be a proper C++ class. You must be able to create many instances of this class. (Please use no third party libraries including C++ STL, Boost etc.)

B. The main program itself. The program must create a circular linked list instance, put new words from the text file to the list or increase their counts there. Also it must be able to ask command options from user for the purpose of (a.) printing most popular words (b.) printing the number of words that is encountered, (c.) printing the counter value of a user given word.

2. Submission

You are expected to submit

a. A working copy of your program executable
b. The Visual Studio project directory (config files) and source code
c. A small report on what you did, sample execution and how to build the application and run. Even better, use a screen capture program like Camtasia Studio, to record a sample run of your executable in video, to make our life easy.

The deadline is set 21 December 2011. Demos will be on 24 December class hours. Submit your files as a zip archive to DELETED_ADDRESS@gmail.com. Change the file extension to "zip_" so google will not complain that there are executables in the file.

The project is 2 PERSON size. The projects made with 3 or more persons are not accepted. 1 person project is not recommended.

3. Cheating Policy.

You are not supposed to use each other's source code. Also please do not use source code from internet, another person or your book's examples. You can be inspired (no direct copy please) from my example singly linked list implementation I had put on OIS.

You are supposed to show pointer operations and please use no helper classes from third party libraries such as STL.

All the source codes will be filtered through a similarity analysis tool, which is known to be effective against many types of code copying and changing tricks.

4. Evaluation

Any lack of 3 items mentioned above (source code with visual studio config files, executable and project report), crashing executables and existence of viruses may cause you to get very low grades. In the demonstrations all the group members must be present. All the members will be asked questions and a common grade will be given to all group members.

I may use another X.txt file to test your program.

5. Bonuses

You can get bonuses for extra efforts :

* Good coding styles and OO programming skills

* Copy constructor, destructor, assignment operator overloading for the circular linked list class.

* Making the circular linked list a generic class (using templates)

* Making the sort function parameterizable with a generic comparison class. (Hence you can sort your list according to any criterion)

* Or any other nice feature you can think of.

Please mention such extra efforts.