Assg 03: Process a File of Scientific Series of Data

CSci 515 Spring 2015

2015-01-30

Dates:

Due: Tuesday February 10, by Midnight

Objectives

- Be able to open a serial text file for reading and writing.

- Process more complex data items from a file of delimiter separated values.

- More practice with I/O stream formatting manipulators.

- Gain more practice using C control structures for implementing algorithms.

- Implement formula calculations into a typical data processing task written in C.

Description

In this assignment, we will be writing a filter, that will filter a file of delimiter separated values of data gathered from an experiment, perform some simple data analysis on the data, and save the results to a new file.

The input file you will be given has the following format (truncated, the real files will have many more items than 7 to process):

```
STRANGE
trial           x               y               z       class
00001    4.23169         4.68996         -8.86438        STRANGE
```

```
00002    5.43040        -3.59577        -4.71896        UP
00003    0.37792         3.34626        -0.22265        CHARM
00004    5.35208         3.96738         1.77813        UP
00005    0.90207        -0.38525         4.66088        CHARM
00006   -4.67474        -4.18064         1.65754        UP
00007    4.27666         4.56251        -8.56897        STRANGE
...
```

The first line represents a filter class upon which the data is to be filtered. More on this below.

The next line after the filter is a header for the columns/features of the data. After the header are the actual data trials in the experiment. Column 1 is simply the trial number for the row of data. Columns 2, 3 and 4 represent x, y and z measured positions of a physical component in some experiment. Column 5 represents a feature category, and will be a string.

Your task is to process this file. We are only interested in experimental trials that are of a particular class. In the above example, the first line indicates that we need to process the **STRANGE** TRIALS. **NOTE**: I can and will use your code on different input files, that will filter on a different class, so you need to read in this first line from the file, and only process subsequent experimental trials of that type.

Each trial of the experiment has recorded the x, y and z position. We need to calculate the distance traveled between successive trials of the target filter class, using the standard euclidean distance:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)}$$

So for example, there are only 2 trials (1 and 7) in our example data of the target **STRANGE** class. The distance traveled for the **STRANGE** item would be:

$$d = \sqrt{(4.27666 - 4.23169)^2 + (4.56251 - 4.68996)^2 + (-8.56897 - -8.86438)^2} = D$$

Assignment Submission

An eCollege dropbox has been created for this assignment. You should upload your version of the out of class assignment by the end of Tuesday 2/3 (midnight) to the dropbox named `Assg 03 Scientific Data File`. Late submissions will not be graded.

Requirements

Your programs must conform to the style and formatting guidelines given for this course. The following is a list of the guidelines that are required for the assignment to be submitted this week.

- The file header and function header for your main function must be present, and filled out correctly.

- You must indent your code correctly and have no embedded tabs in your source code. (Don't forget about the Visual Studio Format Selection command).

- You must not have any statements that are hacks in order to keep your terminal from closing when your program exits.

- You must have a single space before and after each binary operator.

- You must have a single blank line after the end of your declaration of variables at the top of a function, before the first code statement.

Failure to conform to any of these formatting and programming practice guidelines for this assignment will result in a grade of 0 for the assignment, and your program being returned with an indication of which of these items your program violates. Failure to follow other class/textbook programming guidelines may result in a loss of points, especially for those good programming practices given in chapters 1-5 of our textbook which you should have read by now.