COMP 2611 – Data Structures

2020/2021 Semester 1

Guidelines on Group Assignment

Issue with MergeSort

I recently discovered that I made an error in the mergesort algorithm that was supplied in *SortAlgorithms.cpp*. The code given is the version I was working on to prepare the assignment (but not tested fully). It was my intention that your initial experiments would use the version of the mergesort algorithm that we previously discussed in the course (for example, in Lab 8). The code from Lab 8 (*SortMethods.cpp*) is posted on the Group Assignment section of myeLearning. If you do not use this version of mergesort in your experiments, you will not appreciate why I started making the enhancements given in the version you got. Once you observe the difference in performance, you will be able to give a good response to Section 3.

How Many Data Sets?

Some persons have been asking about the size of the data sets that should be used when generating the graphs. The size of the data sets that you use (e.g., 10000, 100000, 1000000) is not important. I would like to you to push each sort algorithm to its limits (when it starts to take > 3 minutes to complete). Whenever that occurs, you should decrease the size of the data set a little at a time to see when it goes from "quite good" to "quite bad". These are the data sets that you want to talk about in the report. For those same data sets, you now want to see what happens with the other algorithms so a comparison can be made. For example, you can say that at 100 million, algorithm *X* is still performing at *y* minutes, but algorithm *A* at *y* minutes could only sort 1 million and is now effectively "out of the competition".

Effort Required

The programming effort required for this assignment is less than, say, Assignment 2. However, the reason I wanted you to work in groups is because I believe the quality of the experiments conducted and the analyses performed will significantly improve by collaboration among three persons. You should view yourselves as scientists conducting a study looking to find answers about which algorithms perform best and why. I am hoping that each group will submit a report that highlights clearly the difference in performance between the algorithms and gives well-thought out explanations for the results obtained. Do not underestimate the importance of submitting a high-quality academic document.

P. Mohan
November 24, 2020