CS 201, Fall 2018

Homework 5

DUE: December 28, Friday @23:55

Description: In this second part of the assignment, you will conduct an experiment with the programs that you have written for Homework 1, 3, and 4. In case you have completed all these assignments, you must have implemented **four different algorithms** that finds the kth largest number among a set of N numbers. You must have implemented **at least two** of these algorithms to obtain partial points from this assignment. The algorithm implementations offer different solutions for the same problem. The goal of the experiment is to compare these solutions with respect to their time efficiency.

1) The Experiment

You will run each of the algorithms for linearly increasing input sizes, i.e., *N*=100000, 200000, 300000, ..., 1000000. The input parameter *k* will be set as *N/2* in all the experiments. You will keep a record of the calculated completion times for each run. If your computer cannot handle input size in the order of millions of numbers, you can select a different linearly increasing input size set such as *N*=100000, 150000, 200000, 250000,..., 500000.

You might observe inconsistent running times especially when *N* is small. So, you should repeat the experiment with a number of randomly generated input sets and take the **average running time** among all. For instance, in order to try solution 1 for *N= 100000*, you can generate 10 different random sets of numbers of size 100000, run the solution for each of these sets, and take the average among all to calculate the average completion time for *N = 100000*. To be able to generate random inputs of different size, you can make use of the **TestInputGenerator** program that was supplied together with Homework 1.

2) The Report

The report must be composed of 3 sections: *(i)* Experimental Setup *(ii)* Results, and *(iii)* Discussion. In the first section (**Experimental setup**), you must describe the data set you used (which input sizes were used? e.g. N = 100000, 15000, 20000, ... etc. or N = 1000, 2000, ... etc.), the algorithms you compared (have you implemented and compared all the four algorithms or a subset of these?) and how you performed the measurements (Did you run each algorithm on each input once? Or did you repeat a number of times and take the average of measurements?).

In the second section (**Results**), you must present the results of your measurements. Hereby, you are expected to include a **table** that lists all the obtained results and create

an **Excel chart** that shows the performance of different algorithms with varying input sizes. Since solution 3 and 4 are extremely more efficient than solutions 1 and 2, the difference between solutions 3 and 4 might not be visible on the chart that compares all the solutions. Hence, you should include an extra chart that compares just solutions 3 and 4 with each other.

In the last section (**Discussion**), you have to interpret and discuss the results. Have you observed the expected outcome in terms of asymptotic complexity calculations regarding algorithms? For instance, do your measurements support the expectation that the running time of the first two algorithms tend to increase exponentially, while the input size increases linearly? Could you finish all the runs? Or have you experienced cases where an algorithm did not even terminate after hours of running?

3) Submission

You will submit this homework via the LMS system. You should follow the file-naming conventions and guidelines below.

- If you did not complete Homework 1, 3, 4 before, and if you complete these to be able to complete this homework, you have to submit a **zip** file containing both your source files and your report in **PDF** format.
- If you completed the previous Homework assignments necessary to complete this one, then you should submit your report as a **PDF** file (No Word, Excel, or other types of files accepted). The name of the file should be in format "**<USER-ID>_hw5.PDF**". For example, if your username is vy1043, then the name of the submitted file should be "vy1043_hw5.PDF". Pay attention that all the letters are in lower-case.
- Late submissions are **not** accepted.
- You can resubmit your homework (until the deadline) if you need to.