Homework 2
CSE 2046 Analysis of Algorithms, Spring 2020
Due 03.05.2020 Sunday 23:59

Design an experiment to compare different algorithms for finding **median** of an unsorted list of n numbers. There are following candidate methods:

(i) Sort all elements by Insertion-sort and return the $\lceil n/2 \rceil$'th element in the list,

(ii) Sort all elements by Merge-sort and return the $\lceil n/2 \rceil$'th element in the list,

(iii) Store all elements in a max-heap and apply $\lfloor n/2 \rfloor$ times max removal. Return the max element in the root,

(iv) Not sort the list, but apply quick select algorithm, which is based on array partitioning, as described in the class. While partitioning, choose the pivot element as the first element in an array,

(v) Apply quick select algorithm, but this time use median-of-three pivot selection[1],

(vi - For groups[3]) Apply quick select algorithm with median-of-medians pivot selection (see https://en.wikipedia.org/wiki/Median_of_medians.)

In this experiment, you are asked to compare these five (or six) methods for finding the median in various input lists (of various sizes and various characteristics). You have to decide on sample inputs to best clarify the performance characteristics of algorithms. While doing your emprical analysis, you may use physical unit of time, count actual number of basic operation's executions, or both. You are also expected to compare your findings with theoretical complexity values. You should provide extensive comments on your results.

This homework has three main steps:

Step 1: Designing the Experiment: This step includes:

(a) Deciding on reasonable inputs / generating reasonable sample inputs[2].

(b) Deciding on reasonable metrics (for complexity measurement).

You should clearly describe all these decisions and reasons behind your decisions in your detailed report.

Step 2: Coding and Running: All algorithms should be implemented in any programming language and experiments should be performed for decided input lists, and results should be evaluated in terms of decided metrics.

Step 3: Illustrating and Analyzing Results: This step includes:

(a) Providing some plots or tables to illustrate performance of the algorithms.

(b) Comparing the performance of different algorithms for different inputs with various sizes.

(c) Describe whether your findings meet theoretical expectations.

You should provide detailed comments for your findings in these comparisons.

Each of the above three steps will be graded separately.

----

[1] **Note 1:** In median-of-three partitioning, the pivot item is selected as the median of the first element, the last element, and the middle element.

[2] **Note 2:** Inputs should reflect best and worst cases. Also you need to use enough (as many as possible) sample inputs to find a reasonable average case time complexity. Also use different inputs with different sizes.

[3] **Note 3:** You may do this homework in groups of two. However, if you perform the project in groups of two, you need to do experiment for all six algorithms (for individual projects, algorithm (vi) is not mandatory). Also you need to clearly declare the division of labor (who did what) in your report. In order to have consistent results (if your performance metric is physical amount of time), it is recommended to run all algorithms in a single computer.
Note 4: Read chapter 2.6 in order to learn guidelines for performing a timing experiment.
Note 5: Input generation time should not be included in the timing experiment.
Note 6: Please submit your commented source codes, input files and detailed report in a zip file that includes both your name(s) and surname by email to **cse246submit@gmail.com**.
Note 7: Do not forget that your grade will mostly determined by quality of your REPORT!