

Homework 1 –Algorithms Fall 2023. (Due October 1st)

Submit clear and well explained answers. Submit one answer per group, but have all the group members names on it. You can use the internet (but not pay wall sites like Chegg) to look for answers, but the final answer must be in your own words. You can also use code you found online for the experiments. Please cite all resources used for answers as well as codes. If you used ChatGPT or other LLMs please mention that.

1. Compare the behavior of a binary min-heap, an AVL tree, and a splay tree, based on the following operations. $6 \times 10^5 = 65$

Perform each operation (Building Search and Insertion) 5 times, and report the individual times as well as the mean and standard deviation

- a. Building: Build each data structure with 50K unique elements. To do this list the numbers from 1 to 50,000. Then use a permutation function to permute the set. Then insert the numbers one by one to build the data structure.(10)
 - i. Report the time taken to build each data set.(10)
 - ii. Report the number of swaps in the heap. The number of rotations in the AVL tree and the Splay tree.(10)
- b. Search:
 - i. Report times to find the 50 lowest numbers in the data structures (10)
 - ii. Report times to find 50 random numbers in the data structures (10)
- c. Insertion
 - i. Insert a set of 5000 random numbers that are not in the original data structure, and report the time for each insertion.(10)

Discuss the behavior of the data structures along with which type of problems are most suitable for which data structures. Use the runtimes to support your analysis. (5).

2. A census is collecting the heights of children from N different schools. Every school sends a sorted list of heights to the census board. The data analyst at the board has to combine these different lists to find the median heights of children in age group 6-8. (35)
 - a. Describe an algorithm in pseudocode, with comments how to find the median (10)
 - b. Given there are N schools with $x_1, x_2, x_3, \dots, x_n$ students respectively, what is the complexity of your algorithm (5)
 - c. Implement your algorithm where given N lists sorted numbers of varying sizes you can find the median of the combination of all the lists. Provide documentation on how your algorithm works (5).
 - d. We will run your algorithms through our input.
 - i. If runtime is within top 10% for code in the same language → 15 points
 - ii. If runtime is within top 50% for code in the same language → 10 points
 - iii. If runtime is within top 100% for code in the same language → 5 points
 - iv. If code does not run -> 0 points.