

CS 4050/7050

Homework 1 (20 points)

This is a programming assignment. Submit all your results in one pdf file.

1. (6 points) Implement the binary search algorithm introduced in the class. You may use any programming language in your implementation. Submit your actual code.
 - a. Input: filename, number.
“filename” is a text file containing a sorted list of distinct integers. “number” is the integer that your program should find its index in the integer list.
 - b. Output: low pointer and high pointer in the first 4 iterations of the algorithm, index, all_time, algo_time,
“index” is the index of the input number, or -1 if the number is not found.
“all_time” is the execution time in milliseconds of your whole program, including file reading time.
“algo_time” is the execution time in milliseconds of the binary search algorithm part of your program, excluding file reading time.
2. (8 points) Run your program on the following 12 test cases. Submit all outputs.
 - a. A list of 100 numbers, data1.txt.: (data1.txt, 19), (data1.txt, 225), (data1.txt, 705)
 - b. A list of 1,000 numbers, data2.txt.: (data2.txt, 128), (data2.txt, 5756), (data2.txt, 9982)
 - c. A list of 10,000 numbers, data3.txt.: (data3.txt, 1997), (data3.txt, 20680), (data3.txt, 23887)
 - d. A list of 100,000 numbers, data4.txt.: (data4.txt, 68189), (data4.txt, 921111), (data4.txt, 935099)
3. (6 points) Perform [least squares linear regression](#) on the two sets of execution times, “all_time” and “algo_time”. You may use existing software or write your own code. Submit two figures, one for each set of times. Plot scatter plot and linear regression line in each figure.
 - a. The figure for “all_time”: the x-axis is $\log_{10}(\text{input size})$, e.g., 2, 3, 4, 5, while the y-axis is the “all_time” execution time in milliseconds (values outputted by your program). The figure should have 12 points (dots) and one straight line (the linear regression function).
 - b. The figure for “algo_time”: the x-axis is $\log_{10}(\text{input size})$, e.g., 2, 3, 4, 5, while the y-axis is the “algo_time” execution time in milliseconds (values outputted by your program). The figure should have 12 points (dots) and one straight line (the linear regression function).