

CIS*3490 The Analysis and Design of Algorithms

Winter 2022
Instructor: Fangju Wang

Assignment 3 (100%)

For this assignment, you are required to write and submit programs. No pseudocode.

Question 1 (40%)

You have a list of n open intervals (a_1, b_1) , (a_2, b_2) , ..., (a_n, b_n) on the integer line. An open interval (a, b) comprises all the points (integers) strictly between its endpoints a and b , i.e., $(a, b) = \{x | a < x < b\}$. (The points do not include a and b .) Find the maximum number of these intervals that have a common point. For example, for the intervals $(0, 5)$, $(-1, 3)$, $(-2, 3)$, $(3, 6)$, this maximum number is 3. The 3 intervals $(0, 5)$, $(-1, 3)$, $(-2, 3)$ have common point 1 or 2.

Write the following programs for finding the maximum number of the intervals that have a common point:

1.1 A program to implement a brute force algorithm. (15%)

1.2 A more efficient program based on the presorting technique. (25%)

When a program is executed, it prompts for a file name, reads in the file containing intervals, and finds the maximum number of the intervals that have a common point. Finally the program reports the number and a common point.

A program is required to report the running time for searching, i.e. finding the number, (in 1.1 for searching, and in 1.2 for sorting and searching). The running time should not include the time for reading the file.

You can use file *data_A3_Q1_1.txt* to develop/test your programs. A different data file will be used to grade your programs. The numbers of intervals and data formats of the two data files will be the same.

Question 2 (60%)

Write the following programs for string search:

2.1 A program to implement a brute force algorithm. (10%)

2.2 A program to implement the Horspool's algorithm. (20%)

2.3 A program to implement the Boyer-Moore algorithm. (20%)

The text is in file *data_A3_Q2.txt*, which has 44049 lines of strings. A search pattern includes the 52 upper and lower case letters only. Search is case-sensitive. When a program is executed,

it reads in the text, prompts for a pattern, finds **all** the occurrences of the pattern in the text, and reports the total number of occurrences found. **Don't** remove any symbols (characters) from the text. A program is required to report the number of pattern shifts and running time for each search.

The file *data_A3_Q2.txt* will be used to grade your programs. You may hard-code the file name in your programs.

2.4 Analyze performance of your brute force and Horspool programs. (10%)

The performance parameters are the number of pattern shifts and running time. To compare two programs, choose ten search patterns of different lengths, and search them by using the programs separately. For each pattern, calculate the ratios of the performance parameters of the two programs. Then, for all the ten patterns, calculate average ratios, and compare and analyze the performance of the two programs in terms of the ratios. Very briefly write your comparison and analysis in the **readme** file submitted with your programs.

Note: Write your own code for this assignment. NO code from any source is allowed.

Due time: 08:00am, Monday March 14, 2022. Submit your work as a tar file to Moodle.