York University EECS 4101 July 7, 2023

Homework Assignment #7
Due: July 14, 2023 at 10:00 p.m.

Reminder: you can use any data structure or algorithm that was described in the textbook or in lectures without having to describe how it works.

1. You are given a collection S of n line segments on the plane $\mathbb{R}^2$. For each segment, one endpoint lies on the line $y = 0$ and the other lies on the line $y = 1$. A segment from $(a, 0)$ to $(b, 1)$ can be described by a pair of numbers, a and b. The entire set S can be described by n such pairs, $(a_1, b_1), \ldots, (a_n, b_n)$. No two segments have a common endpoint. Assume that the segments are sorted so that $a_1 < a_2 < a_3 < \cdots < a_n$.

You must compute how many pairs of segments intersect. Note that any number of segments might intersect in the same point.

[1] **(a)** State a simple condition involving a, b, a', b' that is true iff the segment represented by the pair (a, b) intersects the segment represented by the pair (a', b').

[1] **(b)** Explain how you would determine which of the first $i - 1$ segments intersect with the ith one.

[4] **(c)** Give an algorithm that solves the problem in $O(n \log n)$ time. (Give pseudocode or a description in English, not actual code.) Clearly describe any data structure(s) you use. Your algorithm should be very simple. Briefly explain why your answer to (c) is correct.

Hint: Insert the segments into a data structure one-by-one, and use part (b) after each insertion.

2. Suppose you want to store a set S of jobs. Each job has a priority and a duration. Durations are greater than 0. Assume all the jobs have distinct priorities. You want to support the following operations.

- INSERT(p, d) adds a job with priority p and duration d to S. (Assume no job in S has priority p.)
- EXTRACTMAX removes the job with highest priority from S.
- NUMBERJOBS(D) calculates how many of the highest priority jobs can be done in time D. More precisely, it finds the number of jobs in the largest subset S' of S such that the sum of the durations of jobs in S' is at most D and all priorities of jobs in S' are greater than priorities of jobs in $S - S'$.

We want to use an augmented red-black tree to implement this abstract data type so that every operation runs in $O(\log n)$ time when $|S| = n$.

[1] **(a)** What would you use as the key for ordering the red-black tree?

[2] **(b)** What additional field(s) would you add to each node of the red-black tree?

[2] **(c)** How can you compute the value of your new field(s) of a node x in constant time using information in x and its children?

[2] **(d)** *Briefly* describe how to do an INSERT and EXTRACTMAX.

[3] **(e)** Describe how to do a NUMBERJOBS query. First give an intuitive description, and then give pseudocode.