CPS815/CP8201 - Assignment 1

The only accepted submission format is PDF. It is preferred that you typeset the solutions in LaTeX. Do not overexplain your answers. Do not write code in any specific programming language, only use pseudocode.

1. (40 marks) An open interval, or simply an interval, of the real line $\mathbb{R}$ is denoted by (a, b) where $a, b \in \mathbb{R}$ and $a < b$. A covering set for an interval $(a, b) \subset \mathbb{R}$ is a set of intervals $C = \{(a_1, a_1), \ldots, (a_n, b_n)\}$ such that $(a, b) \subseteq (a_1, b_1) \cup \cdots \cup (a_n, b_n)$. The integer n is called the size of the covering set. A subset of intervals $B \subset C$ that is a covering for $X = (a, b)$ is a called a subcover of C. Given any covering set C for X, we can always find a subcover of C of minimal size. For example, let $X = (17, 21.3)$ and $C = \{(14, 17.1), (15, 18), (12.3, 20), (16.2, 20.7), (19, 22), (18, 21)\}$. Then C is a cover for X. A minimal subcover of C is $B = \{(12.3, 20), (19, 22)\}$.

   Write a greedy algorithm for finding a minimal subcover. The inputs to the algorithm are:

   - (a, b): an interval.
   - $C[1 \ldots n]$: an array of intervals that represents a covering set for (a, b). The interval at index i is $(C[i].a, C[i].b)$.

   Prove that your greedy algorithm is optimal, i.e., it always finds a minimal subcover.

2. (60 marks) Let w be a string of brackets { and }. Then w is called balanced if it satisfies the following recursive definition:

   - w is the empty string, or
   - w is of the form { x } where x is a balanced string, or
   - w is of the form xy where x and y are balanced strings.

   For example, the strings w = {}{{{}}{}} and w = {{}{}} are balanced while the strings w = {{}}}{}} and w = {{}{{}} are not.

   (i) Write an algorithm that determines whether a given string of parentheses is balanced. Analyze the complexity of your algorithm.

   (ii) Write a greedy algorithm to computes the length of the largest balanced substring of a string of parentheses. Prove that your algorithm always outputs the correct answer.

   The complexity of both of your algorithms should be linear in the length of the input string.