CMPT 706 — Algorithms for Big Data

Homework Assignment 4

Instructor: Igor Shinkar *Due date: Tuesday, March 17, 2020*

Submit your solutions, printed or written in readable handwriting, to the assignment boxes in CSIL.

Question 1 (20 points) *Write a linear time algorithm that gets a graph $G = (V, E)$ with n vertices and $n + 10$ edges with weights on its edges, and returns a minimum spanning tree of G. You may assume that all the edge weights are distinct.*

Question 2 (20 points) *Let $G = (V, E)$ be an undirected graph with costs $c_e \geq 0$ on the edges $e \in E$. Assume you are given a minimum cost spanning tree T of G. Now assume that a new edge is added to G, connecting two nodes $v, w \in V$ with cost c.*

(a) Give an efficient algorithm to test if T remains the minimum cost spanning tree with the new edge added to G (but not to the tree T). Make your algorithm run in time $O(|E|)$. Can you do it in $O(|V|)$ time.

(b) Suppose T is no longer a minimum cost spanning tree. Give a linear time algorithm (time $O(|E|)$ to update the tree T to a new minimum cost spanning tree.

Question 3 (20 points) *Prove that for any weighted undirected graph such that the weights are distinct, the minimal spanning tree is unique.*

Question 4 (20 points) *Given a connected undirected graph $G = (V, E)$ with positive weights on the edges. By use of Dijkstra's algorithm, we can find a tree rooted at a vertex s that contains shortest paths from s to any vertex.*

(a) Give an example of a weighted graph, whose minimum spanning tree differs from all its Dijkstra trees.

(b) Prove that a minimum spanning tree and a Dijkstra tree of G always have at least one edge in common.

Question 5 (20 points) *Run the Huffman algorithm on the alphabet $\{A, B, C, D, E, F, G\}$ with frequencies $f_A = 10, f_B = 60, f_C = 10, f_D = 15, f_E = 15, f_F = 40, f_G = 100$. What will be the total length of the encoding of a string over this alphabet with these frequencies.*