University of Toronto
Scarborough Campus
October 16, 2009

CSC C73 Midterm Examination
Instructor: Vassos Hadzilacos

Aids allowed: One 8.5×11 'cheat sheet' (may be written on both sides)

Duration: One hour and fifty minutes

- There should be 7 pages in this exam booklet, including this cover page.
- Answer all questions.
- Put all answers in this booklet, in the spaces provided.
- For rough work, use the backs of the pages; ***these will not be marked***.
- Good luck!

Family Name ______________________________ Given Name ______________________

Student Number ____

Problem	Marks Rec'ved	Marks Worth
1.		10
2.		10
3.		20
4.		30
5.		30
TOTAL		100

QUESTION 1. (10 marks)

The following table gives the frequencies of the symbols of an alphabet.

Symbol	Frequency
A	1/20
B	2/20
C	2/20
D	4/20
E	4/20
F	7/20

In the space below, show a tree that Huffman's algorithm could produce for these symbols and frequencies.

ANSWER:

QUESTION 2. (10 marks)

As we have seen, Dijkstra's shortest path algorithm does not work when edges can have negative lengths. In the space below draw a directed weighted graph, some of whose edges have negative length but which has no cycles of negative length, for which Dijkstra's algorithm does not work correctly.

In your graph identify (a) a path that is computed by Dijkstra's algorithm, but which is not a shortest path between its endpoints, and (b) a shortest path between those endpoints

ANSWER:

QUESTION 3. (20 marks)

If G is a graph and V' is a subset of its vertices, the ***subgraph of*** G ***induced by*** V' is the graph we obtain from G if we delete all vertices except those in V' and all edges except those that connect vertices in V'.

Professor N. O'Bright suggests the following divide-and-conquer algorithm to compute the MST of a connected, undirected graph $G = (V, E)$ with edge weights $w(e)$ for each $e \in E$. If G consists of a single node, the MST trivially has no edges. Otherwise, we partition the set of nodes of G arbitrarily into two sets V_1 and V_2 of about the same size; we recursively compute the MSTs of the subgraphs of G induced by V_1 and V_2; and we join these two MSTs by a minimum weight edge that crosses the V_1/V_2 cut. More precisely, the algorithm that N. O'Bright proposes is this:

D&C-MST(G)
  if $|V| = 1$ **then return** $\varnothing$
  else
    let $V_1 \subseteq V$ be such that $|V_1| = \lceil |V|/2 \rceil$, and $V_2 := V - V_1$
    let G_1 and G_2 be the subgraphs of G induced by V_1 and V_2, respectively
    let e be a minimum weight edge connecting a node in V_1 to a node in V_2
    return D&C-MST(G_1) $\cup$ D&C-MST(G_2) $\cup \{e\}$

Is this algorithm correct? If so, prove it; otherwise, provide a counterexample.

ANSWER:

QUESTION 4. (30 marks)

You walk into a bar in Yorkville knowing that exactly one person in the bar is a celebrity. A ***celebrity*** is a person whom everybody in the bar knows and who knows nobody in the bar (except him or herself). You want to find out who is the celebrity; the only means for doing so is by asking questions of the following kind. You walk up to a person A, you point to some other person B, and ask A: "Do you know that person?" All you get back from A is a truthful "yes" or "no" answer.

Describe a divide-and-conquer algorithm to find out the Yorkville bar celebrity by asking only $O(n)$ questions, where n is the number of people in the bar. Give the recurrence that describes the number of questions asked by your algorithm, and appeal to the "Master Theorem" to justify why your algorithm asks only $O(n)$ questions. For simplicity, you may assume that n is a power of 2.

Hint. If the answer to the question: "Does A know B?" is "yes", what can you conclude about A and B? What if the answer is "no"?

ANSWER:

QUESTION 5. (30 marks)

You are given an array $A[1..n]$ that specifies the locations of n houses on a straight rural road. The location of a house is the distance of that house from the start of the road.

Canada Post must install rural postal "superboxes" along the road to serve these houses. Each superbox can serve multiple houses, but all houses served by a superbox must be within distance d (on either side) of the superbox.

A set of locations (for the superboxes) along the road is ***feasible*** if every house is within distance d of at least one location in the set. A set of locations is ***optimal*** if it is feasible and there is no feasible set of locations with fewer elements. Canada Post is clearly interested in finding an optimal set of locations for the superboxes.

a. (15 marks) Describe an efficient greedy algorithm to determine an optimal set of locations, given A and d.

Hint. Where would you put the superbox that serves the house closest to the start of the road so that it can also serve as many other houses as possible?

ANSWER:

[continued on next page]

b. (15 marks) Prove that the set of locations produced by your algorithm is optimal.

ANSWER:

THE END