College of Computer & Information Science — Spring 2011
Northeastern University — Handout 29
CS5800: Algorithms — 25 April 2011

Final Exam

Instructions.

- **Problems and Points:** There are seven problems, for a total of 60 points. Read all the problems through first and attack them in the order that allows you to make the most progress.
- **Presentation of solutions:** While describing an algorithm, you may use any of the algorithms covered in class or in the text as a subroutine, without elaboration.
- **Proofs:** You need to give proofs only in problems where they are explicitly required.
- **A note on grading:** Your grade on any problem will be determined on the basis of the correctness of your solution and the clarity of the description. In case you are not able to present an algorithm that has the desired properties, give the best algorithm you have designed. Show your work, as partial credit may be given. If you need extra space, use the blank facing page.
- Best of luck!

Question	Score	Maximum
1		7
2		8
3		10
4		10
5		10
6		10
7		5
Total		60

Name: ____________________________________

Problem 1. (7 points) Heavy-hitters: the top k elements

Given a list L of n elements, the *frequency* of an element x in L is the number of times x occurs in L. For instance, if L is the list $[A, B, B, A, C, C, C, A, B, A, D]$, then the frequencies of A, B, C, and D in L are 4, 3, 3, and 1, respectively.

Give an algorithm that takes as input a list L of n elements and an integer k, and returns k distinct elements of L with the highest frequencies in L. If multiple elements have the same frequency, you may break ties arbitrarily. If there are fewer than k distinct elements in L, then your algorithm should return all of the distinct elements in L.

For instance, if the input to your algorithm is the above list L and $k = 2$, then the output should be either $\{A, B\}$ or $\{A, C\}$.

Analyze the worst-case running time of your algorithm. The more efficient your algorithm is in terms of its worst-case running time, the more credit you will get.

Problem 2. (5 + 5 = 10 points total) Pairing in a dance class

A group of n men and n women are attending a dance class. The instructor wants to pair each man with a woman in such a way that the sum of the absolute value of the height differences between partners is minimized.

(a) Let A be a tallest woman in the class (there could be a tie), and let B be a tallest man. Prove that there is an optimal pairing in which A is paired with B.

(b) Based on the result of part (a), give a polynomial-time greedy algorithm for computing an optimal pairing. Analyze the worst-case running time of your algorithm.

Problem 3 (5 + 5 = 10 points) Short-answer questions on minimum spanning trees and network flows

(a) True or False: For any connected weighted undirected graph G, and any node u of G, if (u, v) is an edge that has minimum weight *among all edges incident on* u, then (u, v) belongs to some minimum spanning tree of G.

If your claim is "True", then give a brief proof; if your claim is "False", then provide a counterexample.

(b) Let G be a network flow graph with source s and sink t, and let P be any directed path in G from s to t. Suppose we increase the capacity of every edge in P by one unit. Indicate which one of the following three statements is true. Briefly justify your answer.

(i) It is always the case that (i.e., for every choice of G and P) the maximum flow from s to t increases by exactly one.

(ii) It may be the case that (i.e., there exist G and P such that) the maximum flow from s to t does not increase.

(iii) It may be the case that (i.e., there exist G and P such that) the maximum flow from s to t increases by more than one unit.

Problem 4. (10 points) All-pairs cheap shortest routes

You have been awarded a contract from the American Automobile Association (AAA) to develop a software that determines shortest paths between all pairs of major cities in the United States. You have been supplied a directed graph G over the cities, in which each edge (i, j) represents a one-way highway from city i to city j. For each highway (i, j), you are given a nonnegative distance d_{ij}. Furthermore, each highway is one of two categories: *freeway* or *tollway.*

AAA has determined that while its customers are interested in finding shortest paths, they are particularly concerned about tolls too. Design a polynomial-time algorithm that takes as input the above directed graph G, and a nonnegative integer k, and determines for each pair (i, j) the length of the shortest path from i to j that uses at most k tollways.

Analyze the worst-case running time of your algorithm.

Problem 5. (10 points) Intercepting encoded communication

You have intercepted some communication – a sequence σ of m bits – that you suspect is between your rivals Alice and Bob. You know that every communication between them is a sequence of codewords drawn from a set S of n binary words that they share. And you know the set S. You would like to determine whether σ is indeed a concatenation of words from S.

Give an algorithm that takes σ and S as input and determines whether σ is a concatenation of words from S. Your algorithm must run in time polynomial in m and n. Analyze the worst-case running time of your algorithm.

Example: Suppose S is the set $\{1, 101, 111, 00, 100, 10, 110\}$. If σ is 10111100, then your algorithm should return "yes" since σ can be written as 101+111+00. Note that there are multiple ways that σ can be written as a concatenation of words in S; for instance, 10+111+100 and 10+1+1+1+1+00.

If σ is 00000, then your algorithm should return "no" since σ cannot be written as any concatenation of the words in S.

[Blank page, in case you need more space for Problem 5]

Problem 6. (10 points) Scheduling doctors at a hospital

A hospital is attempting to schedule their doctors for night shifts over the next n nights. There are d doctors. Each doctor specifies a set of nights they are available for the night shift; let S_i denote the set of nights that the ith doctor is available, $1 \leq i \leq d$. The hospital has determined its exact need for each night; for the jth night, $1 \leq j \leq n$, the hospital needs *exactly* p_j doctors to do the night shift. There is one further constraint. To be fair, the hospital does not want any doctor to do more than k night shifts.

Give a polynomial time algorithm to determine an assignment of the d doctors to the n nights so that all relevant constraints are met: every night that the ith doctor is assigned to is in S_i; exactly p_j doctors are assigned on night j; and no doctor is assigned to more than k nights. If no such assignment is possible, then your algorithm should indicate so. (The running time of your algorithm should be polynomial in n and d.)

[Blank page, in case you need more space for Problem 6]

Problem 7. (5 points) Zero-weight directed cycle

Given a directed graph $G = (V, E)$ with integer weights on edges that could be zero, negative or positive, the zero-weight directed cycle problem is to determine whether there exists a directed cycle with weight exactly zero. Show that this problem is NP-complete.

(*Hint:* Use the Directed Hamiltonian Path Problem, or the Directed Rudrata Path Problem for your reduction.)