Spring 2008 **CMSC 498K: Homework 1** Samir Khuller

Due in class: Feb 21.

If you cannot come up with algorithms that run in the required time, then provide (correct) slower algorithms for partial credit. Write your answers using *pseudo-code* in the same style as the textbook. These make the algorithm description precise, and easy to read (as opposed to code in C or some other language).

Please also provide a proof of correctness.

(1) Write out a full proof for the lemma (Munro-Paterson paper) giving bounds on the quantities L_{ij} and M_{ij}. Recall that these were bounds on the least and most number of elements that were larger than the j^{th} element from a sample at level i.

(2) Let X and Y be two arrays, each containing n numbers already in sorted order. Give an $O(\log n)$ algorithm to find the median of *all* $2n$ elements in arrays X and Y.

(3) Develop an $O(n)$ time algorithm that given a set S on n distinct numbers and a positive integer k determines the k numbers that are closest to the median element of S.

(4) We toss 6 identical dice. What is the probability that they all show distinct numbers?

(5) Suppose we have n individuals (imagine n to be very large) - each individual makes phone calls over a certain period of time (say a month). This gives rise to a graph in which we have an edge between two individuals if they had a phone conversation in the last month. Design an efficient algorithm to check if there is a subset of three people such that all three spoke to each other within the last month. What is the running time of your algorithm?