Spring 2008 **CMSC 498K: Homework 5** Samir Khuller

Due in class: May 6.

(1) A team of n players gets off a train. Lets assume each player is numbered with a distinct number chosen from 1 to n. The coach suspects that one player is still on the train. The coach stands at the station exit and needs an easy way of checking who is still on the train as each player exits the platform. How can he/she do this using $O(\log n)$ bits of memory? Do not keep track of the precise set of players that have exited the platform!

(2) Suppose you visit a shopping mall in a town of one million people. During a 2 hour visit to the mall, we spot three obviously pregnant women. That seems like a rather rare event. Is it?

(This question has no right or wrong answer. Just try to estimate what is the chance that you would see a pregnant woman at the mall making reasonable assumptions.)

(3) Prove that the MST of a set of points in the plane is a subgraph of the relative neighborhood graph of that set of points.

(4) In the HISTOGRAM construction problem the goal was to represent a set of n points $x_1, x_2, \ldots x_n$ by a collection of k buckets, such that we stored only one value for each bucket. In other words each point x_i was represented by the value h_j if x_i belonged to the j^{th} bucket. Develop a dynamic programming (optimal) solution for the case where the error of each bucket is not the sum of squared errors but simply the maximum deviation from h_j. We then want to minimize the maximum deviation over all buckets. Example: Suppose the data is 1,4,5,17,16,18,20 and $k = 2$, then we would create buckets with $[1, 4, 5]$ and $[17, 16, 18, 17]$. Clearly $h_1 = 3$ and $h_2 = 17$ and the maximum deviation in the first bucket would be 2 and in the second bucket 1, so the overall cost would be 2.