CMPSCI 711 SPRING 2012: HOMEWORK 4
DUE 2:05 PM, APRIL 30TH

Rules: You may work in groups of at most four. Each group should submit (emailed or handed in at the start of class) one set of typed solutions (remember to include the names of everyone in the group). Cite all sources although work that uses published papers or material from related classes will receive no credit.

Question 1. *Let $\mathcal{A}$ be a stream algorithm that returns the median of a* sorted *list of m values in the range $[n]$ with probability $9/10$. If m is not known in advance, prove that $\mathcal{A}$ must use $\Omega(n)$ memory.*

Question 2. *Let $\mathcal{A}$ be a stream algorithm that determines if a graph on n nodes is bipartite with probability $9/10$. Prove that $\mathcal{A}$ must use $\Omega(n)$ memory. Design a deterministic algorithm that uses $O(n \log n)$ memory (it does not need to support deletions).*

Question 3. *Alice and Bob have strings $x, y \in \{0,1\}^n$ and Bob wants to determine whether $x = y$. Suppose Alice sends a message of length m to help Bob.*

(1) *Alice chooses her message deterministically, prove that $m = \Omega(n)$.*
(2) *Design a random protocol in which m is as small as possible. You may assume players observe a public random string. Bob should return the correct answer with probability $1 - \delta$.*

Question 4. *Modify the F_0 algorithm given in class such that instead of estimating the number of non-zero entries, it estimates the number of odd frequencies.*

Question 5. *In the geometric minimum matching problem, we are given $2n$ points $P = \{p_1, \ldots, p_{2n}\}$. Consider a permutation $f : [2n] \to [2n]$ with no fixed points. We say the cost of permutation f is*

$$\sum_i d(p_i, p_{f(i)})/2 \ .$$

The goal is to approximate the minimum cost over all permutations.

Suppose each $p_i \in \{1, \ldots, \Delta\}^2$ and consider the balanced quaternary tree T whose Δ^2 leaves correspond to all points of the form $\{1, \ldots, \Delta\}^2$. An edge in T between a node in level i and its parent has weight 2^i (we say leaves have level 0). Then define $d(p, q)$ to be the distance in T between the leaves that correspond to p and q. Given a point set P and a node v, let $n(v)$ be the number of points in P that are descendants of v.

(1) *Express the minimum cost of a permutation in terms of the $n(v)$ values. Prove your result.*
(2) *Using the algorithm from the previous question and a probabilistic embedding, design an $O(\log \Delta)$-approximation algorithm for the geometric minimum matching problem when d is the l_1 norm.*