Randomized Algorithms – Final
Fall 2004
Assigned November 23, **Due December 7**

Do any 5 of the following 6 problems. Indicate which problem you are omitting on the top of the first page. I would prefer that you work alone, but will allow you to collaborate *some*. Please indicate on the top of your paper with whom you worked and on which problems. Try to give your partner hints and not complete answers. Exams should be slipped under my office door by the due date. No late papers.

1. You are given a stream of elements that are rapidly flying past you, and you have storage enough to hold only k elements. The goal is to maintain a random sample of k elements from the stream. I.e., when you have seen t of the stream elements, the storage must contain a random one of the $\binom{t}{k}$ possible k-subsets; moreover, this should hold for all times $t \geq k$. How would you do this? (Note that if you do not put an element into storage, it is gone forever.)

2. Prove that if $p \geq m$ is a prime, then
$$\mathcal{H} = \{h_{a,b}(x) = ((ax + b) \ (\mathrm{mod} \ p) \ (\mathrm{mod} \ n) | a, b \in \mathbf{Z}_p, a \neq 0\}$$
   is a 2-universal hash family from $M \to N$, with $|M| = m$ and $|N| = n$. If we remove the restriction $a \neq 0$, does the family still remain 2-universal?

3. Imagine we are using randomized rounding to solve the $s - t$ MinCut problem. First, we formulate the problem as an integer linear program. Given a cut that separates s and t, let $x_i = 0$ for all vertices i that are reachable from s, and $x_i = 1$ for all other vertices. (Thus $x_s = 0$ and $x_t = 1$.) Let $z_{i,j}$ be the indicator variable for edge $\{i, j\}$ crossing the cut, so $z_{i,j} = |x_i - x_j|$. Thus, we get the following:
$$\text{minimize} \sum_{\{i,j\} \in E} c_{i,j} z_{i,j} \text{ subject to}$$
$$\forall \{i, j\} \in E, z_{i,j} \geq (x_j - x_i);$$
$$x_s = 0 \text{ and } x_t = 1;$$
$$\forall i, \forall j, x_i, z_{i,j} \in \{0, 1\}.$$

   Show that the optimal solution to the linear relaxation gives an integral solution, and hence rounding is not even necessary. (Hint: Think about how you would use randomized rounding to go from the optimal solution to the linear program to an integer solution. Then argue that there must have been an integral optimal solution as well.)

1

4. A (d, c, α)-expander is a graph $G = (V, E)$ where each vertex has degree at most d and every subset $S \subseteq V$ with at most cn vertices has $|N(S)| \geq \alpha|S|$. ($N(S)$ is the set of neighbors of S and may include some nodes in S.)

Starting with a set V of n vertices, add a random matching between the vertices as follows: (a) choose a random permutation $v_1, v_2, ..., v_n$ of the vertices and (b) add the edges (i, v_i) for all i. (We may have parallel edges and self-loops.) Perform this process $d = 600$ times.

Prove that $G = (V, E)$ is a $(2d, \frac{7}{20}, \frac{3}{2})$-expander with probability at least $1/2$. (Hint: what is the probability that some set S with $|S| \leq cn$ does not expand?)

5. Show that the cover time of a random walk on a connected d-regular graph is $O(n^2 \log n)$. (Note that this quantity is independent of d.)

6. In this question you will prove that random walks on constant degree expanders mix very rapidly: for any subset of vertices C that is fairly large, the probability that a random walk of length ℓ avoids C is $\exp(-\ell)$. Let $G = (V, E)$ be an unweighted undirected graph and A be its adjacency matrix.

a) Sow that the number of walks of length ℓ is $g^T A^\ell g$ where g is the all-1 vector.

b) Suppose now that G is d-regular and has n vertices. Suppose that each of the eigenvalues of A except the largest (which is d) has magnitude at most λ. Let C be a subset of cn vertices and let A' be the adjacency matrix of the induced graph on $V \setminus C$. Show that every eigenvalue of A' is at most $(1 - c)d + c\lambda$ in magnitude.

c) Conclude that if $\lambda < 0.9d$ (i.e., G is an expander) and $c = 1/2$ then the probability that a random walk of length ℓ in G (starting at a randomly chosen vertex) avoids C is at most $\exp(-\ell)$.