6.842 Randomness and Computation	April 28, 2008

Homework 5

Lecturer: Ronitt Rubinfeld — *Due Date: May 7, 2008*

1. Prove that for every function Ext : $\{0,1\}^n \rightarrow \{0,1\}$, there is an SV-source X with parameter δ such that $Pr[\text{Ext}(X) = 1] \leq \delta$ or $Pr[\text{Ext}(X) = 1] \geq 1 - \delta$.

2. First prove a seemingly strange claim about extractors that take semi-random bits but get no truly random bits, and then use it to show something interesting about extractors that take both semi-random and truly random bits. (As defined in class, if the Ext function has only one parameter, then it refers to the former type of extractor, otherwise it refers to the latter type).

   (a) Show the following: For any n and any $k < n$ and any flat k-source X, if an extractor Ext : $\{0,1\}^n \rightarrow \{0,1\}^m$ with $m = k - 2\log\frac{1}{\epsilon} - O(1)$ is chosen at random from the functions mapping $\{0,1\}^n$ to $\{0,1\}^m$, then with probability at least $1 - 2^{-\Omega(2^k \epsilon^2)}$, $\text{Ext}(X)$ is ϵ-close to U_m (the uniform distribution on m bits) with respect to statistical distance.

   (b) Then show that: For any n and $k < n$, and any $\epsilon > 0$, there exists a (k, ϵ)-extractor Ext : $\{0,1\}^n \times \{0,1\}^d \rightarrow \{0,1\}^m$ with $m = k + d - 2\log\frac{1}{\epsilon} - O(1)$ and $d = \log(n - k) + 2\log\frac{1}{\epsilon} + O(1)$.

3. We say that an undirected graph on n nodes is *labeled* if the edges adjacent to each vertex are labeled with numbers from 1 to n, and no two edges are labeled with the same number. An edge may be labeled differently on each of its endpoints.

   Given is a labeled graph G on n nodes, a node v in G, and a string $s = (s_1, \ldots, s_k) \in \{1, \ldots, n\}^\star$. Consider the following procedure. Our initial position is v. In the i-th step, if there is an edge adjacent to the current node, labeled with s_i, we follow that edge. Otherwise, we stay at the current node. We call s a (G, v)-*cover* if it can be used to visit all vertices of G by following to the above procedure.

   A string $s \in \{1, \ldots, n\}^\star$ is a *universal traversal sequence* for size n if for every labeled connected graph G on n nodes and every node v in G, s is a (G, v)-cover.

   (a) Show that there exists a universal traversal sequence for size n of length $n^{O(1)}$.

   (b) Show that there exists a universal traversal sequence for size n of length $n^{O(\log n)}$ that can be constructed in $n^{O(\log n)}$ time.

   Hint: You may use the following outline:

   - Design a logarithmic space algorithm $\mathcal{A}$ that given a labeled graph G and two nodes i and j, simulates a random walk starting from i, and accepts the input if it visits j, which happens with probability at least $1 - \frac{1}{3n^2}$, if i and j are in the same connected component.
   - Derandomize $\mathcal{A}$ so that it still runs in polynomial time, but uses only $O((\log n)^2)$ random bits, and accepts every input with i and j in the same connected component with probability at least $1 - \frac{1}{2n^2}$. Denote the new algorihm by $\mathcal{A}'$.

- Argue that for each labeled connected graph G on n nodes there is a seed r such that for each pair of nodes (i, j), $\mathcal{A}'$ accepts (G, i, j).
- Conclude the proof of the claim, by showing how to construct a universal traversal sequence.

4. (optional, don't turn it in) Recall that a bipartite graph (V_1, V_2, E) is a (K, α) *bipartite expander*, if every set $W \subseteq V_1$ of size at most K is adjacent to at least $\alpha|W|$ vertices in V_2. In this problem, we consider a $(K, (1-\epsilon)D)$ bipartite expander $G = (V_1, V_2, E)$ with the degree of each vertex in V_1 equal to D. Prove first the following properties of G:

   - For a set $W \subseteq V_1$, a vertex $v \in V_2$ is a *unique neighbor* of W if it is incident to exactly one edge from W. Show that every set $W \subseteq V_1$ of size at most K has at least $(1-2\epsilon)D|W|$ unique neighbors.
   - Show that for every set $W \subseteq V_1$ of size at most $K/2$, there are at most $|W|/2$ vertices in $V_1 \backslash W$ that share at least δD neighbors with W, for some $\delta = O(\epsilon)$.

   Let $N = |V_1|$ and $M = |V_2|$. We will now see how to use G to store a small subset Y (where $|Y| \leq K/2$) of a large domain X (where $|X| = N$), and test membership in Y by probing only one bit. We identify elements of X with nodes in V_1. From now on, assume that $X = V_1$. Our data structure consists of M bits assigned to nodes of V_2. Ideally, we would like that for any element $x \in X$, all the bits assigned to the neighbors of x be 1, if $x \in Y$, and 0, otherwise. We cannot achieve this property in small space, but instead we can get the following relaxed property.

   > **Property P:** For all $x \in X$, all but a $\delta = O(\epsilon)$ fraction of the neighbors of x are assigned 1, if $x \in Y$, and 0, if $x \notin Y$.

   Prove the following:

   - If we store an assignment that has Property P, we can test membership in Y with error probability δ, by reading one bit of the data structure.
   - An assignment satisfying Property P exists for any Y of size at most $K/2$.

   It turns out that there exist expanders that we need with $M = O(K \log N)$, for any $\epsilon > 0$. Therefore, we can construct a data structure of size $O(K \log N)$ which allows for testing membership in $Y \subseteq \{1, \ldots, N\}$ by probing just one bit.

5. (optional, don't turn it in) Let $\mathcal{A}$ be a one-sided error polynomial-time algorithm for a language $\mathcal{L}$ that errs with probability at most $1/2$, and uses r random bits. We saw in class how to turn $\mathcal{A}$ into a one-sided error polynomial-time algorithm that errs with probability at most 2^{-k}, and only uses $r + O(k)$ random bits. Show that techniques you learnt in class can be used to get a similar algorithm that only uses $O(r \log k)$ random bits.