CMSC 858C, Spring 2009: Homework 2, due on March 12th (Thursday) at the start of class.

Notes: Please work on this with your group-mate(s): each group needs to submit only one solution. Consulting other sources (including the Web) is not allowed. Write your solutions **neatly**; if you are able to make partial progress by making some additional assumptions, then state these assumptions clearly and submit your partial solution.

1. Derandomize Turán's Theorem that we saw in class, on large independent sets in graphs. Give an interpretation as a greedy algorithm. **(5 points)**

2. Suppose we have n balls in an urn, with some balls of color Red, and others of different colors. Suppose we draw balls at random from the bin, *without* replacement. Let E_i be the event that the ith ball drawn is Red. Show that these events are negatively correlated in the following sense: for any distinct indices $i_1, i_2, \ldots, i_k$,

$$\Pr[E_{i_k} \mid E_{i_1}, E_{i,2}, \ldots, E_{i_{k-1}}] \leq \Pr[E_{i_k}].$$

(8 points)

3. Let $\oplus$ denote the usual XOR function. Suppose an integer n and some $\epsilon \in [0, 1/2]$ are given. Prove that there exists a multiset (recall that repetitions are allowed in a multiset) S of n-bit strings with the following two properties: (i) S has cardinality at most $O(n/\epsilon^2)$; (ii) Suppose a vector $X = (X_1, X_2, \ldots, X_n)$ is sampled uniformly at random from S. (Recall that S is a multiset. So, if a string s occurs k times in S, then $\Pr[X = s] = k/|S|$.) Then, for all nonempty $A \subseteq \{1, 2, \ldots, n\}$,

$$1/2 - \epsilon \leq \Pr\left[\left(\bigoplus_{i \in A} X_i\right) = 1\right] \leq 1/2 + \epsilon.$$

(**Hint:** Use the probabilistic method. Use care with the quantification in the question.) **(8 points)**

4. Suppose there is an unknown quantity $v > 0$ that we want to estimate, and are given some value $\epsilon \in (0, 1)$. Suppose also that we have an efficient random process $\mathcal{R}$ to generate a random variable X with the guarantee that $\Pr[(1-\epsilon)v \leq X \leq (1+\epsilon)v] \geq 3/4$. Now suppose we want to boost this confidence of "3/4" to "$1 - \delta$", for some given $\delta < 1/4$; i.e., we want to efficiently generate a random real Y such that $\Pr[(1 - \epsilon)v \leq Y \leq (1 + \epsilon)v] \geq 1 - \delta$. Show how to sample $O(\log(1/\delta))$ times independently from process $\mathcal{R}$ and suitably aggregate these, in order to generate Y. Prove your claim. **(8 points)**

5. Let $X_1, X_2, \ldots, X_n$ be pairwise independent random variables taking values in $\{0, 1\}$, such that $X = \sum_i X_i$ has mean 1. The X_i are not necessarily identically distributed. Prove that $\Pr[X > 0] \geq 1/2$. **(5 points)**