Complexity Homework 4

Released: March 10, 2010

Due: March 31, 2010

Problem 1:

This is a quick refresher for basic probability concepts. A probability distribution over a (finite) set S is a function $\pi : S \to [0,1]$ such that $\sum_{x \in S} \pi(x) = 1$. A (real-valued) random variable X is a function $X : S \to \mathbb{R}$ along with a probability distribution π. We define $\mathbf{Pr}_{s \leftarrow \pi}[X(s) = x] = \sum_{s:X(s)=x} \pi(s)$ (often shortened to $\mathbf{Pr}[X = x]$, when π is understood). We define expectation $\mathbf{E}_{s \leftarrow \pi}[X(s)] = \sum_{s \in S} X(s) \cdot \pi(s)$ (often shortened to $\mathbf{E}[X]$, when π is understood).

(a) (Linearity of expectation.) Given two random variables X_1, X_2, define a new random variable X as $X(s) = aX_1(s) + bX_2(s)$ (for some real numbers a and b). Show that $\mathbf{E}[X(s)] = a\mathbf{E}[X_1(s)] + b\mathbf{E}[X_2(s)]$.

(b) (Markov's inequality.) Given a non-negative random variable X, show that $\mathbf{Pr}[X > t\mu] < 1/t$, where $\mu = \mathbf{E}[X]$. (Here $t > 0$.)

(c) Given a random variable X, suppose we define a new random variable Z_X as $Z_X(s) = X(s) - \mu$ where $\mu = \mathbf{E}[X]$. Calculate $\mathbf{E}[Z_X]$.

(d) (Variance and Chebyshev's inequality.) Given a random variable X, define a new random variable Z_X as $Z_X(s) = (X(s) - \mu)^2$ where $\mu = \mathbf{E}[X]$. Then the variance of X is defined as $\mathbf{Var}(X) = \mathbf{E}[Z_X]$ and the standard deviation as $\sigma(X) = \sqrt{\mathbf{Var}(X)}$. Use Markov's inequality to bound $\mathbf{Pr}[|X - \mu| > t\sigma(X)]$. (This is called Chebyshev's inequality.)

(e) Two random variables X and Y are said to be independent if for all real numbers x, y, $\mathbf{Pr}[X = x \text{ and } Y = y] = \mathbf{Pr}[X = x]\,\mathbf{Pr}[Y = y]$. Show that if X and Y are independent, $\mathbf{Var}(X + Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$. Further, if $\{X_i\}_{i=1}^t$ are t random variables which are *pairwise independent* (that is, X_i and X_j are independent for all $i \neq j$), show that $\mathbf{Var}(\sum_i X_i) = \sum_i \mathbf{Var}(X_i)$.

(f) Suppose $\{X_i\}_{i=1}^t$ are t pairwise independent random variables which take binary (0-1) values such that $\mathbf{Pr}[X_i = 1] = p$ for all i. Use Chebyshev's inequality to prove that

$$\mathbf{Pr}\left[\left|\frac{\sum_{i=1}^t X_i}{t} - p\right| > \delta\right] = O\left(\frac{1}{\delta^2 t}\right).$$

Problem 2:

Let M be a probabilistic TM. Define the *gap* of M for a language L to be $\min_{x \in L} \mathbf{Pr}[M(x) = \texttt{yes}] - \max_{x \notin L} \mathbf{Pr}[M(x) = \texttt{yes}]$. and its *error* for L to be $\max_x \mathbf{Pr}[M(x) \neq L(x)]$. Bound the gap and error in terms of each other.

Problem 3:

Define Expected-Time-**PP** to be the class of languages decided by probabilistic Turing machines (via acceptance probability $> \frac{1}{2}$) whose *expected* running-time is polynomial (as opposed to **PP**, where the running time is worst-case polynomial). Show that **EXP** $\subseteq$ Expected-Time-**PP**. What can you say about inclusion in Expected-Time-**PP** for classes larger than **EXP**? What if the expected running time is restricted to be constant instead of polynomial?

Problem 4:

In this problem we shall prove impossibility of deterministic extraction from Santha-Vazirani sources. We work with probability distributions over $S = \{0,1\}^n$, the set of n-bit strings.

For $x \in \{0,1\}^n$, let x_i denote the i-th bit of x and $x_{\bar{i}}$ denote the other $n-1$ bits of x. Call a distribution π *δ-balanced at position i* if for all $y \in \{0,1\}^{n-1}$, $\mathbf{Pr}[x_i = 0 | x_{\bar{i}} = y]$ and $\mathbf{Pr}[x_i = 1 | x_{\bar{i}} = y]$ differ by at most δ.

(a) Verify that π is δ-balanced at position i if and only if for every $y \in \{0,1\}^{n-1}$,

$$\frac{1-\delta}{1+\delta} \leq \frac{\pi(y_1 \ldots y_{i-1} 0 y_i \ldots y_{n-1})}{\pi(y_1 \ldots y_{i-1} 1 y_i \ldots y_{n-1})} \leq \frac{1+\delta}{1-\delta}.$$

Call a distribution δ-balanced if it is δ-balanced at all positions $i = 1, \ldots, n$. Note that if the output distribution of a randomness source is δ-balanced it is a Santha-Vazirani source (but not vice-versa).

Consider an arbitrary boolean function $f : \{0,1\}^n \to \{0,1\}$. Let π_0^f be the probability that $f(x) = 0$ when x is drawn according to the distribution π. That is, $\pi_0^f = \sum_{x|f(x)=0} \pi(x)$. Similarly let $\pi_1^f = \sum_{x|f(x)=1} \pi(x)$.

(b) Show that for every $f : \{0,1\}^n \to \{0,1\}$, and every $\delta \in [0,1]$, there exists a δ-balanced distribution π over $\{0,1\}^n$ such that $|\pi_0^f - \pi_1^f| \geq \delta$.
(Hint: Consider separately the functions f for which $|\mathcal{U}_0^f - \mathcal{U}_1^f| \geq \delta$ and those for which $|\mathcal{U}_0^f - \mathcal{U}_1^f| < \delta$, where $\mathcal{U}$ is the uniform distribution over n-bit strings.)

Conclude that there are no simple (deterministic) extractors which can extract a single ϵ-balanced bit from all δ-balanced Santha-Vazirani sources, with $\epsilon < \delta$.

Problem 5:

(a) (Randomized rounding.) Given a probability distribution ρ over R and random variable X, with range $[0,1]$, define a probability distribution π over $S = R \times \{0,1\}$ as follows:

$$\text{For } r \in R : \pi((r,1)) = \rho(r) \cdot X(r) \text{ and } \pi((r,0)) = \rho(r)[1 - X(r)]$$

Verify that π is indeed a valid probability distribution. Now define a binary random variable Z (i.e., with range $\{0,1\}$), with underlying probability distribution π, as $Z(r,0) = 0$ and $Z(r,1) = 1$ for all $r \in R$. Show that $\mathbf{E}[Z] = \mathbf{E}[X]$.
(That is, instead of the real variable X, the binary random variable Z can be used without changing the expectation (though the variance could increase). This is called randomized rounding because Z can be considered to be sampled as follows: draw a sample from X, and using the value obtained as the bias, flip a coin, to get a *rounded* (0-1) value.)

(b) (Deterministic rounding.) Let X be as above. Consider a new random variable Z^* defined over R and with respect to the same probability distribution ρ, as follows: $Z^*(r) = 1$ if $X(r) > \frac{1}{2}$ and 0 otherwise. Using Markov's inequality, show that $2\mathbf{E}[X] - 1 \leq \mathbf{Pr}[Z^* = 1] \leq 2\mathbf{E}[X]$. Conclude that if $\mathbf{E}[X] > 7/8$ then $\mathbf{Pr}[Z^* = 1] > 3/4$ and if $\mathbf{E}[X] < 1/8$ then $\mathbf{Pr}[Z^* = 1] < 1/4$.

(c) (Eliminating an auxiliary random source.) In this problem we consider a randomized algorithm A which draws its randomness from two independent random sources, a "main" source (with an arbitrary distribution) and an auxiliary *perfect* random source. Our goal is to change it to an algorithm B which uses only the main source, by enumerating over all random strings from the auxiliary source (while drawing only as many bits as A draws from the main source).
Describe B so that if the probability of error of A is at most 1/8 (when run using the two sources), then the probability of error of B is at most 1/4 (when run using only the main source). Prove that B has these properties. (*Hint: Use part (b). What should the real variable X be?*)

Problem 6 (Extra Credit):

In this problem we use basic linear algebra to analyze (weak) extraction from an SV source (see Lecture 15).

(a) (Collision probability.) Define a probability distribution π over $\{0,1\}^d$. We will view π as a real vector of length 2^d (i.e. $\pi \in \mathbb{R}^{2^d}$), such that (with elements indexed by $i \in \{0,1\}^d$) $\pi_i = \pi(i)$. Define collision probability of π, $\text{col}(\pi)$ to be the probability that two strings drawn independently according to π are the same. Show that $\text{col}(\pi) = \|\pi\|^2$, where $\|v\|$ is defined as $\sqrt{\langle v, v\rangle}$.

(b) (An orthonormal basis.) Define 2^d vectors $\rho^{(s)}$ (for $s \in \{0,1\}^d$) as follows: $\rho_j^{(s)} = \frac{1}{2^{d/2}}(-1)^{\langle s,j\rangle}$. Note that $\|\rho^{(s)}\| = 1$, and each element in $\rho^{(s)}$ is $\pm\frac{1}{2^{d/2}}$, the sign depending on whether $\langle s, j\rangle$ is even or odd. Show that $\langle \rho^{(s)}, \rho^{(t)}\rangle = 0$ for all $s \neq t$.

(*Hint: $s \neq t$ means there is at least one position where the vectors s and t differ. Use this to show that all the vectors can be partitioned into pairs (j_0, j_1) such that the parities of $\langle s, j_0\rangle$ and $\langle t, j_0\rangle$ are equal, and those of $\langle s, j_1\rangle$ and $\langle t, j_1\rangle$ are different.*)

Hence these 2^d vectors form an orthonormal basis for the vector space $\mathbb{R}^{2^d}$. This basis is called the *Fourier Basis.*

(c) (Change of basis.) Recall that given an orthonormal basis any vector v can be written as a linear combination of the basis vectors, with the coefficients being the inner product of the vector v with basis vectors. So we can write $\pi = \sum_s \langle \pi, \rho^{(s)}\rangle \rho^{(s)}$. Use this to rewrite $\|\pi\|^2$.

(d) Consider the extractor which on input $r \in \{0,1\}^d$ and seed $s \in \{0,1\}^d$ outputs the bit $\langle r, s\rangle_2$. (Here we use $\langle \cdot, \cdot\rangle_2$ to denote dot product as integers modulo 2.) We consider feeding the extractor an input drawn according to the distribution π. For each seed value s, define $\mathrm{Gap}_s^\pi = \mathbf{Pr}_{r\leftarrow\pi}[\langle r, s\rangle_2 = 0] - \mathbf{Pr}_{r\leftarrow\pi}[\langle r, s\rangle_2 = 1]$. Show that $\mathrm{Gap}_s^\pi = 2^{d/2}\langle \pi, \rho^{(s)}\rangle$.

(e) Deduce that $\mathbf{E}_{s\leftarrow\mathcal{U}_d}[(\mathrm{Gap}_s^\pi)^2] = \mathrm{col}(\pi)$, where $\mathcal{U}_d$ is the uniform distribution over $\{0,1\}^d$.

(f) From this, using the fact that $\mathbf{E}[X]^2 \leq \mathbf{E}[X^2]$, conclude that

$$|\,\mathbf{Pr}_{r\leftarrow\pi, s\leftarrow\mathcal{U}_d}[\langle r, s\rangle_2 = 0] - \mathbf{Pr}_{r\leftarrow\pi, s\leftarrow\mathcal{U}_d}[\langle r, s\rangle_2 = 1]| \leq \|\pi\|.$$

Note that the left hand side is the bias of the extracted bit, when the input r is drawn according to the distribution π and the seed s is drawn independently from $\mathcal{U}_d$. Finally, show that when π is an SV source with bias bounded by a constant less than 1, $\|\pi\| = 2^{-\Omega(d)}$.