CS 225: Pseudorandomness	**Prof. Salil Vadhan**
Problem Set 1	
Assigned: Tue. Feb. 6, 2007	Due: Wed. Feb. 21, 2007(1 PM)

- Recall that your problem set solutions must be typed. You can email your solutions to `cs225-hw@eecs.harvard.edu`, or turn in it to MD138. You may write formulas or diagrams by hand. Aim for clarity and conciseness in your solutions, emphasizing the main ideas over low-level details.
- If you use LaTeX, please submit both the source (`.tex`) and the compiled file (`.ps`). Name your files `PS1-yourlastname`.
- Starred problems (like Problem 5) are extra credit.

Problem 1. (Robustness of the model) Suppose we modify our model of randomized computation to allow the algorithm to obtain a random element of $\{1, \ldots, m\}$ for any number m whose binary representation it has already computed (as opposed to just allowing it access to random bits). Show that this would not change the classes **BPP** and **RP**.

Problem 2. (A Different Identity Test) In this problem, you will analyze an alternative to the identity testing algorithm seen in class, which directly handles polynomials of degree larger than the field size.

The following definitions and facts may be useful: A polynomial $p(x)$ over a field $\mathbb{F}$ is called *irreducible* if it has no nontrivial factors (i.e. factors other than constants from $\mathbb{F}$ or constant multiples of p). Analogously to prime factorization of integers, every polynomial over $\mathbb{F}$ can be factored into irreducible polynomials and this factorization is unique (up to reordering and constant multiples). It is known that the number of irreducible polynomials of degree at most d over a field $\mathbb{F}$ is at least $\mathbb{F}^{d+1}/2d$. (This is similar to the Prime Number Theorem for integers, but is much easier to prove.) For polynomials $p(x)$ and $q(x)$, $p(x) \bmod q(x)$ is the remainder when p is divided by q. (More background on polynomials over finite fields can be found in the references in the syllabus.)

In this problem, we consider a version of the identity testing problem where a polynomial $p(x_1, \ldots, x_n)$ over finite field $\mathbb{F}$ is presented as a formula built up from elements of $\mathbb{F}$ and the variables $x_1, \ldots, x_n$ using addition, multiplication, and *exponentiation* with exponents given in *binary*. We also assume that we are given a representation of $\mathbb{F}$ enabling addition, multiplication, and division in $\mathbb{F}$ to be done quickly.

1. Let $p(x)$ be a polynomial of degree $\leq D$ over a field $\mathbb{F}$. Prove that if $p(x)$ is nonzero (as a formal polynomial) and $q(x)$ is a randomly selected polynomial of degree at most $d = O(\log D)$, then the probability that $p(x) \bmod q(x)$ is nonzero is at least $\Omega(1/\log D)$. Deduce a randomized, polynomial-time identity test for *univariate* polynomials presented in the above form.
2. Obtain an identity test for multivariate polynomials by reduction to the univariate case.

Problem 3. (Primality Testing)

1. Show that for every positive integer n, $(x+1)^n \equiv x^n + 1 \pmod{n}$ iff n is prime.

2. Obtain a **co-RP** algorithm for PRIMALITY$= \{n : n \text{ prime}\}$ using Part 1 together with the previous problem. (In your analysis, remember that the integers modulo n are a field only when n is prime.)

Problem 4. (A Chernoff Bound) Let $X_1, \ldots, X_t$ be independent $[0,1]$-valued random variables, and $X = \sum_{i=1}^{t} X_i$.

1. Show that for every $r \in [0, 1/2]$, $\mathrm{E}[e^{rX}] \leq e^{r\,\mathrm{E}[X]+r^2 t}$. (Hint: $1 + x \leq e^x \leq 1 + x + x^2$ for all $x \in [0, 1/2]$.)

2. Deduce the following Chernoff Bound: $\Pr\left[X \geq \mathrm{E}[X] + \varepsilon t\right] \leq e^{-\varepsilon^2 t/4}$. Were did you use the independence of the x_i's?

Problem 5. (Necessity of Randomness for Identity Testing*) In this problem, we consider the "oracle version" of the identity testing problem, where an arbitrary polynomial $p : \mathbb{F}^m \to \mathbb{F}$ of degree d is given as an oracle (ie black box) and the problem is to test whether $p \equiv 0$. Show that any deterministic algorithm that solves this problem when $m = d = n$ must make at least 2^n queries to the oracle (in contrast to the randomized identity testing algorithm from class, which makes only one query provided that $|\mathbb{F}| \geq 2n$).

Is this a proof that $\mathbf{P} \neq \mathbf{RP}$?