CS 225 - Pseudorandomness	Prof. Salil Vadhan

Problem Set 1

Harvard SEAS - Spring 2015 — *Due: Fri. Feb. 13, 2015*

Your problem set solutions must be typed (in e.g. LaTeX) and submitted electronically to `cs225-hw@seas.harvard.edu`. You are allowed 12 late days for the semester, of which at most 5 can be used on any individual problem set. (1 late day = 24 hours exactly). Please name your file `PS1-`lastname`.*`.

The problem sets may require a lot of thought, so be sure to start them early. You are encouraged to discuss the course material and the homework problems with each other in small groups (2-3 people). Identify your collaborators on your submission. Discussion of homework problems may include brainstorming and verbally walking through possible solutions, but should not include one person telling the others how to solve the problem. In addition, each person must write up their solutions independently, and these write-ups should not be checked against each other or passed around.

Strive for clarity and conciseness in your solutions, emphasizing the main ideas over low-level details. Do not despair if you cannot solve all the problems! Difficult problems are included to stimulate your thinking and for your enjoyment, not to overwork you. *ed problems are extra credit.

Problem 2.3.(Zero error vs. 1-sided error) Prove that $\mathbf{ZPP} = \mathbf{RP} \cap \mathbf{co\text{-}RP}$.

Problem 2.5.(Identity Testing via Modular Reduction) In this problem, you will analyze an alternative to the algorithm seen in class, which directly handles polynomials of degree larger than the field size. It is based on the same idea as Problem 2.4, using the fact that polynomials over a field have many of the same algebraic properties as the integers.

The following definitions and facts may be useful: A polynomial $f(x)$ 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 f). 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 mentioned in Problem 2.4, but is easier to prove.) For polynomials $f(x)$ and $g(x)$, $f(x) \bmod g(x)$ is the remainder when f is divided by g. (More background on polynomials over finite fields can be found in the references listed in Section 2.6.)

In this problem, we consider a version of the Identity Testing problem where a polynomial $f(x_1, \ldots, x_n)$ 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 $f(x)$ be a univariate polynomial of degree $\leq D$ over a field $\mathbb{F}$. Prove that there are constants c, c' such that if $f(x)$ is nonzero (as a formal polynomial) and $g(x)$ is a randomly selected polynomial of degree at most $d = c \log D$, then the probability that $f(x) \bmod g(x)$

is nonzero is at least $1/c' \log D$. Deduce a randomized, polynomial-time identity test for *univariate* polynomials presented in the above form.

2. Obtain an identity test for multivariate polynomials by (deterministic) reduction to the univariate case.

Problem 2.6.(PRIMALITY)

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

2. Obtain a **co-RP** algorithm for the language PRIMALITY$= \{n : n \text{ prime}\}$ using Part 1 together with Problem 2.5. (In your analysis, remember that the integers modulo n are a field only when n is prime.)

Problem 2.7.(A Chernoff Bound) Let $X_1, \ldots, X_t$ be independent $[0,1]$-valued random variables, and $X = \sum_{i=1}^{t} X_i$. (Note that, in contrast to the statement of Theorem 2.21, here we are writing X for the sum of the X_i's rather than their average.)

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. Deduce the Chernoff Bound of Theorem 2.21: $\Pr\left[X \geq \mathrm{E}[X] + \varepsilon t\right] \leq e^{-\varepsilon^2 t/4}$ and $\Pr\left[X \leq \mathrm{E}[X] - \varepsilon t\right] \leq e^{-\varepsilon^2 t/4}$.

3. Where did you use the independence of the X_i's?

Problem 2.8.(Necessity of Randomness for Identity Testing*) In this problem, we consider the "oracle version" of the identity testing problem, where an arbitrary polynomial $f : \mathbb{F}^m \to \mathbb{F}$ of degree d is given as an oracle and the problem is to test whether $f = 0$. 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 $|\mathbb{F}| \geq 2n$).