Probabilistic Algorithms: Exercise 1

Submission Date: November 11th

January 17, 2005

We define three variants of polynomial time probabilistic algorithms that solve decision problems. (For a language L we use $1_L(x)$ to denote the function that outputs 1 on $x \in L$ and 0 otherwise).

Definition 1 (ZPP). *A language L is in ZPP if there is a polynomial $p(n)$ and a randomized algorithm A such that for every x:*

- $E(Time_A(x)) \leq p(|x|)$.
- $\Pr[A(x) = 1_L(x)] = 1$.

Definition 2 (BPP). *A language L is in BPP_ϵ if there is a polynomial $p(n)$ and a randomized algorithm A such that for every x:*

- $\Pr[Time_A(x) \leq p(|x|)] = 1$.
- $\Pr[A(x) = 1_L(x)] \geq 1 - \epsilon$

We define $BPP = BPP_{1/3}$.

Definition 3 (RP). *A language L is in RP_ϵ if there is a polynomial $p(n)$ and a randomized algorithm A such that for every x:*

- $\Pr[Time_A(x) \leq p(|x|)] = 1$.
- *If $x \in L$ then* $\Pr[A(x) = 1_L(x)] \geq 1 - \epsilon$.
- *If $x \notin L$ then* $\Pr[A(x) = 1_L(x)] = 1$.

We define $RP = RP_{1/3}$.

Theorem 4 (Markov's inequality). *Let X be a non negative random variable.*

$$\Pr[X \geq aE(X)] \leq 1/a$$

Theorem 5 (Chernoff's inequality). *Let $X_1, \ldots, X_t$ be independent 0/1 variables. Let $X = \sum_i X_i$ then for any $0 < \delta < 1$*

$$\Pr[|X - E(X)| \geq \delta E(X)] \leq 2^{-\frac{\delta^2 E(X)}{6}}$$

1. (Relation between classes) Show that:

    (a) $RP \subseteq NP$.

    (b) $ZPP \subseteq RP \subseteq BPP$. (Hint: Use Markov's inequaality).

    (c) $RP = RP_{1-1/|x|} = RP_{1/2^{|x|}}$.

    (d) $BPP = BPP_{1/2-1/|x|} = BPP_{1/2^{|x|}}$. (Hint: Use Chernoff's inequality).

    (e) $ZPP = RP \cap coRP$. (Here $coRP = \{L : coL \in RP\}$ and $coL = \{x : x \notin L\}$.)

2. (Sublinear time for approximately finding the k'th element in an array) In class we saw an algorithm that for any a and t and an (unsorted) array S of length n, the algorithm queries the array t times and with probability $1 - 2^{-\Omega(a^2)}$ finds an element in the array that is at distance at most $an/\sqrt{t}$ from the median of the array. The median is the k'th element in the array for $k = n/2$. Modify this algorithm so that it uses the same resources and works for any k. (Hint: Recall that the algorithm in class sampled t elements at random from the array.)

3. (A probabilistic construction of error correcting codes)

    (a) Let x, y be independent uniformly chosen strings in $\{0,1\}^n$. Show that $\Pr[d_H(x,y) < n/4] \leq 2^{-\Omega(n)}$. (Recall that $d_H(x,y)$ is the number of coordinates in which they differ).

    (b) Show that there exist $\ell = 2^{\Omega(n)}$ strings $x_1, \cdots, x_\ell \in \{0,1\}^n$ such that for each $i \neq j$, $d_H(x,y) \geq n/4$.