Today
- Complexity Classes
- Error Reduction
- Sampling

CS5.413.1
Pseudorandomness
Lecture 04 (2021-9-2)
Instructor: Prahladh Harsha.

Recall from last lecture:
pairwise independence.

Pairwise Independent family of hash functions

A family $\mathcal{H} = \{h : [N] \to [M]\}$ is said to pw ind. family if

(1). $H \leftarrow \mathcal{H}$, $\forall x \in [N]$ $H(x)$ - uniform in U_M.

(2). $\forall x_1 \neq x_2 \in [N]$, $H \leftarrow \mathcal{H}$
$H(x_1)$ & $H(x_2)$ - independent

Equivalently.

$\forall x_1 \neq x_2 \in [N]$, $\forall y_1, y_2 \in [M]$

$$\Pr_{H \leftarrow \mathcal{H}} [H(x_1) = y_1 \wedge H(x_2) = y_2] = \frac{1}{M^2}.$$

Last time: Construction : $N = M = |\mathbb{F}|$

$$\mathcal{H} = \{ h_{a,b} \mid a, b \in \mathbb{F} \}$$
$$h_{a,b} : \mathbb{F} \to \mathbb{F}$$
$$x \mapsto ax + b.$$

Claim: $\mathcal{H}_{a,b}$ is pw ind family.

---

bits needed to specify $H \in \mathcal{H}$.

Independent: $|\mathbb{F}| \log |\mathbb{F}|$

Pairwise Independent: $2 \log |\mathbb{F}|$
(construction from last time)

$N = 2^n$; $M = 2^m$

$h : \{0,1\}^n \to \{0,1\}^m$

Case (i) $n = m$; $\mathbb{F} = GF(2^n)$

Case (ii) $n < m$; $\mathbb{F} = GF(2^m)$

$h : \{0,1\}^n \to \{0,1\}^m$

$h(x) = h'(\underbrace{x\ 0^{m-n}}_{m})$

Case (iii) $n > m$: $\mathbb{F} = GF(2^n)$

$h(x) = h'(x)\big|_m$

bits reqd to generate h
$= 2 \max\{m, n\}$.

In fact, $\max\{m,n\} + m$.

<u>Thm</u>. $\forall\, m,n$, there exists a pw ind family of hash functions $\mathcal{H}_{m,n}$ that requires at most $2\max\{m,n\}$ bits to specify any fn $h \in \mathcal{H}_{m,n}$.

Complexity Classes:

P, BPP, RP, coRP, ZPP, ...

Decision Problems / Languages.
YES/NO problems. | $L \subseteq \{0,1\}^*$ (assuming Boolean alphabet)

Deterministic Algorithms.

$$x \longrightarrow \boxed{A} \longrightarrow A(x)$$ - YES/NO
acc/rej

$t : \mathbb{N} \to \mathbb{N}$.

Algorithm A runs in time t if $\forall$ inputs x,
A runs in time at most $t(|x|)$.

t - n^2, n^3, n^c, 2^n, 2^{n^2}, ...

P: Set of languages that have a poly time deterministic alg A s.t.

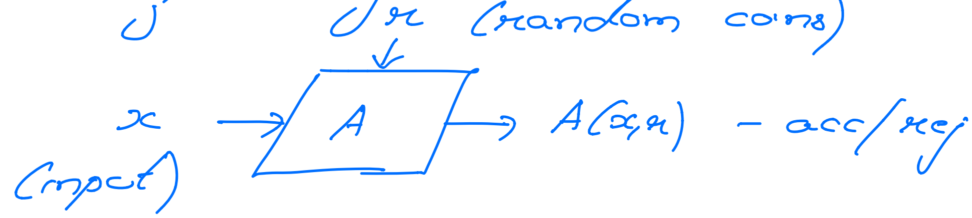
$$\begin{cases} x \in L \Rightarrow A(x) = acc \\ x \notin L \Rightarrow A(x) = rej \end{cases}$$

$A : \{0,1\}^* \to \{acc, rej\}$

$\cup|$

L

Randomized Algorithms



Algorithms A - that run in a fixed time for all inputs of a particular length (irrespective of random coins)

RP: (randomized polynomial time).

- Set of languages L for which there exists a randomized poly time Algorithm A. satisfying

$$x \in L \Rightarrow \Pr_r[A(x,r) = acc] \geq 1/2.$$
$$x \notin L \Rightarrow \Pr_r[A(x,r) = acc] = 0$$

<u>coRP</u>: Same as above except
$$x \in L \Rightarrow \Pr_r[A(x,r) = acc] = 1$$
$$x \notin L \Rightarrow \Pr_r[A(x,r) = acc] \leq 1/2.$$

eg: PRIMES $\in$ coRP
Polynomial Identity Testing

Error Reduction for RP:

$A^{(t)}$: On input x
- Pick random $r_1 \ldots r_t$
- Run $A(x,r_1), \ldots A(x,r_2) \quad A(x,r_t)$
- Acc if any one of them acc
  & rej otherwise.

$$x \in L \Rightarrow \Pr_{R=r_1..r_t}[A^{(t)}(x,\bar{R}) = acc] \geq 1-\left(\frac{1}{2}\right)^t$$
$$x \notin L \Rightarrow \Pr_R[A^{(t)}(x,\bar{R}) = acc] = 0$$

Algorithms w/ err on both sides

BPP : $L \in BPP$

if there is a poly time rand alg A s.t.

$$x \in L \Rightarrow \Pr_r[A(x,r) = acc] \geq 3/4$$

$$x \notin L \Rightarrow \Pr_r[A(x,r) = acc] \leq 1/4.$$

error $= 1/4$

$$\begin{array}{ccc} & BPP & \\ & \cup\!\mid & \\ \cup\!\mid & & \supseteq \\ coRP & & RP \\ \supseteq & & \cup\!\mid \\ & ZPP & \\ & \cup\!\mid & \\ & P. & \end{array}$$

ZPP: (zero error. probabilistic poly time).

run-time – randomized

output – correct

$L \in ZPP$. if there is a randomized poly time alg A that runs in expected polynomial time:

$$x \in L \Rightarrow \Pr_r[A(x,r) = acc] = 1$$

$$x \notin L \Rightarrow \Pr_r[A(x,r) = acc] = 0.$$

Prop: (1) $ZPP = RP \cap coRP$.

(2) $ZPP \subseteq BPP$

Error Reduction for BPP:

Basic Probability Inequalities / Tail Bounds.

(1) **Markov's Inequality**: X is a non-negative random variable w/ finite expectation $\mathbb{E}[X]$.

then $\Pr[X \geq \alpha] \leq \mathbb{E}[X]/\alpha$.

(2) **Chebyshev's Inequality**:

X is real r.v w/ finite exp $\mathbb{E}[X]$ & finite variance $\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

then
$$\Pr[|X - \mathbb{E}X| \geq \varepsilon] \leq \frac{\mathrm{Var}[X]}{\varepsilon^2}$$

(3) **Chernoff Bound**:

$X_1 \ldots X_t$ - t independent $[0,1]$-valued random variables,

$$\bar{X} = \frac{\sum X_i}{t} \quad ; \quad \mu = \mathbb{E}[\bar{X}]$$

$$\Pr[|\bar{X}-\mu| > \varepsilon] \leq 2e^{-t\varepsilon^2/4}$$

Next: Chernoff Bound to reduce error in BPP

Thm: The following 3 stmts are equivalent:

(1) $L \in BPP$ (ie error $\leq 1/4$)

(2) $\forall$ polynomial $p(n)$, L has a rand (two-sided error) polytime alg w/ error $\leq \frac{1}{2^{p(n)}}$

(3) $\exists$ poly $q(n)$, L has a rand poly time alg w/ error $\leq \frac{1}{2} - \frac{1}{q(n)}$

Pf: (2) $\Rightarrow$ (1) $\Rightarrow$ (3). - easy.

(3) $\Rightarrow$ (2) : Let A be the rand alg from (3).

Construct $A^{(t)}$: On input x

1. Pick $r_1 \ldots r_t$ - ind rand coins

2. Run $A(x, r_1) \ldots$

$\cdot\ A(x, r_t)$

3. Acc if majority accepts.

- Analysis: error of $A^{(t)}$,

Fix $x \in \{0,1\}^n$, X_i - 0/1- valued random variable.

$$X_i = \begin{cases} 1 & \text{if } A(x, R_i) \text{ is incorrect} \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbb{E}[X_i] = \Pr[A(x, R_i) \text{ is incorrect}]$$

$$\leq \frac{1}{2} - \frac{1}{q(n)}$$

$$\Pr[A^{(t)} \text{ is incorrect on } x]$$

$$= \Pr\left[\sum X_i \geq t/2\right]$$

$$= \Pr\left[\frac{\sum X_i}{t} \geq \frac{1}{2}\right]$$

$$\leq \Pr\left[\frac{\sum X_i}{t} - \left(\frac{1}{2} - \frac{1}{q(n)}\right) \geq \frac{1}{q(n)}\right]$$

$$\leq \Pr\left[|\bar{X} - \mu| \geq \frac{1}{q(n)}\right]$$

$$\leq 2\exp\left(-\frac{t}{4q^2(n)}\right)$$

$\leq \frac{1}{2^{p(n)}}$ if $t = C p(n) q^2(n)$

Running time of $A^{(t)}$ = Running time $A * t$

---

t- Repetition of Algorithm A.

\# random bits $(A^{(t)})$

$= t \cdot$ \# random bits (A)

Qn: What if $R_1 \ldots R_t$ - were only pairwise independent?

Tail bound for sum of pairwise ind r.v:

Let $X_1 \ldots X_t$ - pairwise ind $[0,1]$-val. random variables. $\bar{X} = \frac{1}{t} \sum X_i$.

$\mu = \mathbb{E}[\bar{X}]$.

$\Pr[|\bar{X} - \mu| > \varepsilon] \leq$

Pf: $\mathrm{Var}(\bar{X}) = \mathrm{Var}\left(\frac{\sum X_i}{t}\right)$

$$= \mathbb{E}\left[\left(\frac{\sum X_i}{t}\right)^2\right] - \left(\mathbb{E}\left[\frac{\sum X_i}{t}\right]\right)^2$$

$$= \frac{1}{t^2}\left[\mathbb{E}\left[\sum_{i,j} X_i X_j\right] - \left(\sum \mathbb{E}[X_i]\right)^2\right]$$

$$= \frac{1}{t^2}\left[\sum_i \mathbb{E}[X_i^2] + 2\sum_{i<j} \mathbb{E}[X_i X_j] - \sum_{i,j} \mathbb{E}[X_i]\,\mathbb{E}[X_j]\right]$$

$$= \frac{1}{t^2}\left[\sum_i \left(\mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2\right) + \sum_{i\neq j}\left(\mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\,\mathbb{E}[X_j]\right)\right]$$

$$= \frac{1}{t^2}\left[\sum_i \left(\mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2\right)\right]$$

(pairwise independence)

$$= \frac{1}{t^2}\sum \mathrm{Var}(X_i)$$

$$\Pr\left[|\bar{X} - \mu| \geq \varepsilon\right] \leq \frac{\mathrm{Var}[\bar{X}]}{\varepsilon^2}$$

$$= \frac{1}{t^2}\,\frac{\sum \mathrm{Var}(X_i)}{\varepsilon^2}$$

$$\leq \frac{\sum 1}{t^2 \varepsilon^2} \quad \left(\text{since } \mathrm{Var}(X_i) \leq 1\right)$$

$$= \frac{1}{t\varepsilon^2}.$$

In Chernoff, to reduce error to δ we set $t \leftarrow O\left(\frac{1}{\varepsilon^2} \log \frac{1}{\delta}\right)$.

But w/ pairwise independence, in order to reduce error to δ we need to set

$$t \leftarrow \frac{1}{\varepsilon^2 \delta}$$