Pseudorandomness - Lecture 17

Date:
Instructor: Ramprasad
Lecture: #17

Agenda: - PRGs from hardness.
[Nisan-Wigderson] thms.

Recap: - $G : \{0,1\}^d \to \{0,1\}^m$ is an ε-PRG for a class C if $\forall$ $A \in C$,

$$\left| \mathop{\mathbb{E}}_{x \sim U_m}[A(x)] - \mathop{\mathbb{E}}_{y \sim U_d}[A(G(y))] \right| \leq \varepsilon.$$

- PRG $\Leftrightarrow$ Next bit unpredictability.

  Given $x_1, \ldots, x_{i-1}$, can you predict the i^{th} bit with some advantage?

- [Nisan] - PRG for logspace with seed length $d = O(\log^2 n)$.

- Ideal seed length for C = size s circuits:
  $$d = O(\log s + \log 1/\varepsilon).$$

- If we find an explicit PRG with above s.l, then BPP = P.

Revisiting NBU:

Defn: $G : \{0,1\}^d \to \{0,1\}^m$ is ε-next bit predictable by a class C if $\exists$ $i \in [m]$ and $A \in C$ such that.

$$\Pr_{x = G(U_d)}[A(x_1, \ldots, x_{i-1}) = x_i] \geq \frac{1}{2} + \varepsilon.$$

Lemma: G is ε-PRG for size s $\Rightarrow$ G is ε-NBU for size s.
And, G is ε-NBU for size s $\Rightarrow$ G is $m\varepsilon$-PRG for size $s - O(1)$.
(Saw in lecture 13).

How do we build PRGs?

▷ Toy case: Getting stretch 1. ie $d \to d+1$

$$G(x_1, \ldots, x_d) = x_1, \ldots, x_d, g(x_1, \ldots, x_d)$$

for some $g : \{0,1\}^d \to \{0,1\}$

What do we want from g?

For any size s circuit C, we want.

$$\Pr_{x \sim \mathcal{U}_d}\left[C(x_1, \ldots, x_d) = g(x_1, \ldots, x_d) \right] \leq \frac{1}{2} + \varepsilon$$

ε-hard to even guess $g(x)$ on a randomly chosen x. "g is ε-hard to guess by size s ckts"

$\therefore$ If we can find a "hard" function g, then we have a PRG that stretches by 1 bit. Meh.

More stretch:

$$G : \{0,1\}^{dk} \to \{0,1\}^{dk+k}.$$

$$x^{(1)}, \ldots, x^{(k)} \mapsto x^{(1)}, \ldots, x^{(k)}; g(x^{(1)}), \ldots, g(x^{(k)}).$$

Same argument will show that G is NBU.

Now stretches by k bits ... still Meh.

How do we get more stretch ?

why?!

If g is acting on disjoint subset of seed bits, then $\frac{\text{output length}}{\text{input length}} \leq \frac{d+1}{d}$

Idea: What if g acts on "almost disjoint" subsets?

Defn: (Combinatorial designs): A collection of subsets $S_1, \ldots, S_m \subseteq [d]$ is an (l, a) design if they satisfy the following properties.

- $|S_i| = l$
- $|S_i \cap S_j| < a$ whenever $i \neq j$.

We want a "large" collection of small-ish sets with very small pairwise intersection.

Do such designs exist?

Let's try to pick random sets one at a time.

$S_1, \ldots, S_t$ already picked. S_{t+1} chosen at random.

$\Pr[S_{t+1} \text{ is valid}] = \Pr\left[|S_{t+1} \cap S_i| < a \quad \forall i \in [t]\right]$

$\therefore \Pr[S_{t+1} \text{ is } \underline{\text{not}} \text{ valid}] \leq \sum_i \Pr[|S \cap S_i| \geq a]$

$$\Pr\left[|S \cap S_i| \geq a\right] \leq \frac{\binom{l}{a} \cdot \binom{d}{l-a}}{\binom{d}{l}} \qquad (*)$$

$$\frac{\binom{d}{l-a}}{\binom{d}{l}} = \frac{\cancel{d!}}{(l-a)!\,(d-l+a)!} \cdot \frac{l!\,(d-l)!}{\cancel{d!}} = \frac{l!}{(l-a)!} \cdot \frac{a!}{a!} \cdot \frac{(d-l)!}{(d-l+a)!}$$

$$= \binom{l}{a} \Big/ \binom{d-l+a}{a} \leq \binom{l}{a} \Big/ \binom{d-l}{a}$$

$$\therefore \Pr\left[|S_i \cap S| \geq a\right] \leq \frac{\binom{l}{a}^2}{\binom{d-l}{a}}$$

$\therefore$ As long as $t < \binom{d-l}{a} \Big/ \binom{l}{a}^2$, we have

$$\Pr\left[|S \cap S_i| \geq a \text{ for some } i\right] < 1.$$

$\Rightarrow$ There exists a good S to continue.

$\therefore$ There exist (l,a) designs involving $\binom{d-l}{a} \Big/ \binom{l}{a}^2$ sets.

$l = 10\log n$. $\qquad a = \log n \qquad d = 1000\log n$

$$\binom{d-l}{a} \approx n^{11.4} \qquad\qquad \binom{l}{a} \approx n^{10 H(1/10)}.$$

$$\approx n^{4.6}$$

$\Rightarrow$ We have $\geq n^2$ such sets!

More generally, for any $l \geq 2a$, we can find comb. designs $S_1, \ldots, S_m \subseteq [d]$ with $m = 2^a$ and $d = 2l^2/a$.

Are there explicit constructions? Yes! (PSET 3 ;-))

What do we do with this?

Suppose $h : \{0,1\}^l \to \{0,1\}$ is a "hard to guess" function, then here is a candidate generator:

$$G : \{0,1\}^d \longrightarrow \{0,1\}^m.$$

$$G(z_1, \ldots, z_d) = \left(h(z|_{S_1}), \ldots, h(z|_{S_m}) \right)$$

Thm: Suppose C is a circuit of size $\leq s$ that [NW] ε-next bit predicts G, then we can build a circuit C' that $\frac{1}{2}+\varepsilon$ approximates h, with $\text{size}(C') \leq s + m \cdot 2^a$.

In other words, if h is hard enough, then G is indeed a PRG.

Eg: Suppose $\{ h_r : \{0,1\}^r \to \{0,1\} \}$ is ε-hard to guess for size $s(r) = 2^{r/100}$

$G : \{0,1\}^d \longrightarrow \{0,1\}^m$, NW generator with $a = \log m$, $l = 100 \log(m^3)$, $d = \square \log m$. with h_l used inside.

If G is Next-bit-predictable by size m ckts, then says h_ℓ is ε-guessable by size $\leq m + a \cdot 2^a \cdot \leq m^3$

$2^{\ell/100}$

$\Rightarrow\!\!\Leftarrow$.

What is the seed length? $O(\log n)$

How much time does $G(y)$ take? $\text{Time}(h) \cdot m$

$= \text{poly}(m)$ if h is computable in $2^{O(\ell)}$ time.

$\therefore$ BPP = P.

Hardness of h_r		Consequence
$2^{\Omega(r)}$	$\leftrightarrow$	BPP = P
2^{r^ε}	$\leftrightarrow$	$\text{BPP} \subseteq \text{Time}(2^{\text{polylog}(n)})$
$r^{\omega(1)}$	$\leftrightarrow$	$\text{BPP} \subseteq \bigcap_{\varepsilon > 0} \text{TIME}(2^{n^\varepsilon})$

Proof of the NW Theorem:

Thm: Suppose C is a circuit of size $\leq s$ that ε-next bit predicts G, then we can build a circuit C' that $\frac{1}{2}+\varepsilon$ approximates h, with $\text{size}(C') \leq s + m \cdot 2^a$.

Pf: Say C is a Next-bit predictor. Then, $\exists i \in [m]$ such that

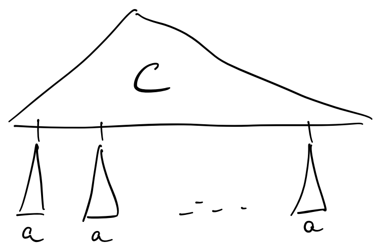
$$\Pr_{z \sim \mathcal{U}_d} \left[C\left(h(z|_{S_1}), .., h(z|_{S_{i-1}})\right) = h(z|_{S_i}) \right] \geq \frac{1}{2} + \varepsilon.$$

$$\Pr_{\text{blue}} \left[C(h(z|_{S_1}), .., h(z|_{S_{i-1}})) = h(y) \right] \geq \frac{1}{2} + \varepsilon.$$

$$\Pr_{\text{blue}} \left[C\left(h^{(1)}(y^{(1)}), .., h^{(i-1)}(y^{(i-1)})\right) = h(y) \right] \geq \frac{1}{2} + \varepsilon.$$

Obs: Each $y^{(i)}$ is only 'a' variables!

$\therefore$ Each $h^{(j)}(y^{(j)})$ can be computed by a stupid circuit of size 2^a.



$= C'$.

$$\Pr\left[C'(y) = h(y) \right] \geq \frac{1}{2} + \varepsilon.$$

and $\text{size}(C') \leq \text{size}(C) + m \cdot 2^a$. $\square$

Avg case hardness:

Any circuit C of size $\leq s$,

$$\Pr_x\left[C(x) = h(x) \right] \leq \frac{1}{2} + \varepsilon$$

Any small circuit makes lots of mistakes.

Worst-case hardness:

Any circuit C of size s makes <u>some</u> mistake.

ie $\exists x$: $C(x) \neq h(x)$.

[Impagliazzo-Wigderson]: Converting a worst case hard function into an avg case hard function to plug into NW.

Replace all the avg case hyp. to worst case hyp.

Next class: An exposition of the IW transformation.
(without proofs)