Pseudorandomness - Lecture 18.

Instructor: Ramprasad
Date: 2021-10-28.

Agenda: - PRGs from weaker assumptions (an exposition).

Recap: ▷ $G: \{0,1\}^d \to \{0,1\}^m$ for a class $\mathcal{C}$. (typically, size m ckts).

$\forall A \in \mathcal{C}$

$$\left| \underset{x \sim \mathcal{U}_m}{\mathbb{E}}[A(x)] - \underset{y \sim \mathcal{U}_d}{\mathbb{E}}[A(G(y))] \right| \leq \varepsilon.$$

▷ An explicit PRG $\Rightarrow$ derandomisation in $\approx \mathrm{poly}(2^d, m)$ time.

▷ $\{h_r: \{0,1\}^r \to \{0,1\}\}$ is ε-hard to guess (ε-strongly hard.) for size $s(r)$ if for any circuit $C_r: \{0,1\}^r \to \{0,1\}$ of size $\leq s(r)$,

$$\Pr_{x \in \mathcal{U}_r}[C_r(x) = h_r(x)] \leq \frac{1}{2} + \varepsilon$$

▷ Thm [Nisan-Wigderson]: If there is a family of fns $\{h_r: \{0,1\}^r \to \{0,1\}\}$ that is computable in $2^{O(r)}$ time, that is hard-to-guess for size $s(r)$, then there is an explicit PRG for size m ckts

$$G: \{0,1\}^{d(m)} \to \{0,1\}^m$$

with $$d(m) = O\left(\frac{1}{\underbrace{\log m}_{a}} \cdot \big(\underbrace{s^{-1}(m^3)}_{\ell}\big)^2\right)$$

If $s(r) = 2^{r/100}$ $\Rightarrow$ $s^{-1}(m^3) = 100 \log(m^3)$

$\Rightarrow$ $d = O(\log m)$

$s(r) = 2^{\sqrt{r}}$ $\Rightarrow$ $s^{-1}(m^3) = (\log m^3)^2$

$\Rightarrow d = O((\log m)^3)$

How can we strengthen this theorem?

Weakening hypothesis

Strengthening conclusion.

Weaker Hypotheses

ε-hard to guess $\approx \frac{1}{2}-\varepsilon$ weakly hard

Defn: (ε-weakly hard): A fn $f: \{0,1\}^m \to \{0,1\}$ is ε-weakly hard ~~to guess~~ for size s circuits if $\forall$ ckts C of size s,

$$\Pr_x [f(x) = C(x)] \leq 1-\varepsilon.$$

"Any size s circuits makes ε-fraction of errors"

Qn: Suppose I give you a weakly hard-to-guess function, can you build a strongly hard to guess one?

Qn: Suppose you have a coin that is biased... $\Pr[\text{heads}] = 0.9$.
How do you create a "close to unbiased" bit?

Toss k times. And output the parity of heads.

Ex: $\Pr[\text{odd heads}] = \frac{1}{2} - \frac{1}{2}(2p-1)^k$

Yao's XOR Lemma: Suppose $f: \{0,1\}^n \to \{0,1\}$ is δ-weakly hard for size s. Then $f^{\oplus k}: \{0,1\}^{nk} \to \{0,1\}$, given by

$f^{\oplus k}(x^{(1)},\dots,x^{(k)}) = f(x_1) \oplus \cdots \oplus f(x_k)$
is $\varepsilon + (1-\delta)^k$ - strongly hard to guess for size $s' = \Theta(\varepsilon^2 \delta^2 s)$

∴ Finding a weakly hard fn is sufficient as we can "boost" the hardness.

$h : \{0,1\}^n \to \{0,1\}^k$ $\qquad \Pr_x[C(x) = h(x)] \le 1-\delta$

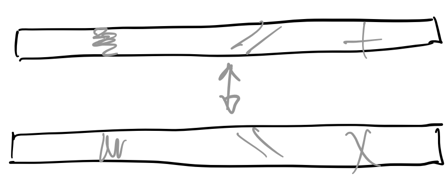
Might be easier to find multi-output hard fns.
$f : \{0,1\}^m \to \{0,1\}^k$.

Lemma [Goldreich-Levin] Let $g : \{0,1\}^m \times \{0,1\}^k \to \{0,1\}$
given by $g(x,r) = \langle r, f(x) \rangle \bmod 2$.
Then, if f is weakly hard, then g is strongly hard.

∴ Finding a weakly hard multi-output fn is sufficient.

A different perspective wrt approx computation.

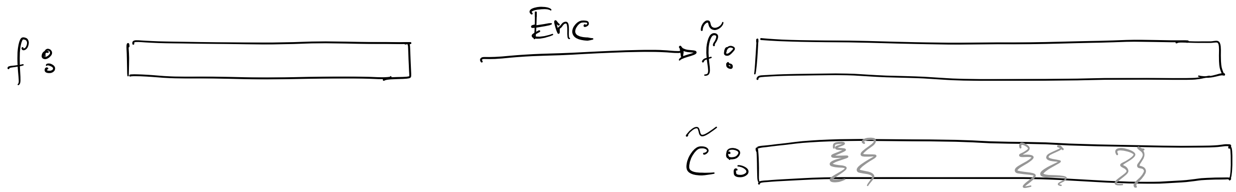
$f : \{0,1\}^m \to \{0,1\}$



$\{0,1\}^{2^n}$.

$C : \{0,1\}^m \to \{0,1\}$
that approx. f 90%. → these two vectors differ in ≤10% of coords.
$\Pr[C(x) = f(x)] \ge 90\%$

Can we "amplify" a single mistake into "many" disagreements? Smell like an error correcting code.



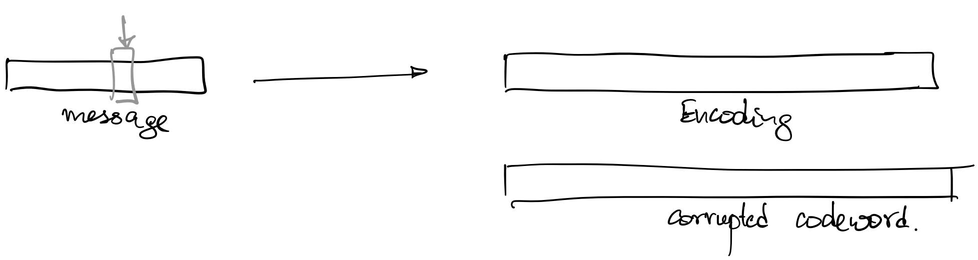
↳ agrees with $\tilde{f}$ on 90% of coords $\equiv$ received word.

Then, if the code is efficiently decodable, then we can "decode" f using $\tilde{C}$. ... not so fast.

Suppose $\tilde{f} = \text{Enc}(f)$, and $\tilde{C}$ is a size s circuit that satisfies $\Pr[\tilde{C}(y) = \tilde{f}(y).] \geq 1-\varepsilon$.
We want to build a circuit C (not too large) that uses $\tilde{C}$ and computes f correctly everywhere.

Input to C : $x \in \{0,1\}^m$.

Want to use the TT of $\tilde{C}$ sparingly



Any message bit can be recovered using "few" queries on the corrupted codeword.

Locally decodable codes.

Eg: Hadamard code: $\mathrm{Had} : \{0,1\}^k \to \{0,1\}^{2^k}$

$$\mathrm{Had}(v) = \left(\langle v, x \rangle \;:\; x \in \{0,1\}^k \right)$$

$\mathrm{Had}(101) =$

000	001	010	011	100	101	110	111
0	1						

Claim: This code is locally decodable from $\varepsilon = 0.1$ errors using $O(1)$ queries.

Pf: Algo:

Want i^{th} message bit.

Pick $y \in_R \{0,1\}^k$.

Query location y to get a $= \langle x, y \rangle$ w.p 90%

Query location $y + e_i$ to get b. $= \langle x, y+e_i \rangle$ w.p 90%

Return $a + b \bmod 2$.

$= \langle x, e_i \rangle$ w.p $\geq 80\%$.

$\square$.

$\therefore$ We do have locally decodable codes.. but Had has horrible rate.

But Reed-Muller codes have all the props we want.

$(\mathrm{RM} \circ \mathrm{Had})$

Lem: There are codes $Enc: \{0,1\}^L \to \{0,1\}^{L'}$ with $L' = poly(L)$ that locally decodable from $1/50$ errors. using just $polylog(L)$ many queries.

Cor: Suppose $f: \{0,1\}^m \to \{0,1\}$ is any fn, and say $Enc(f) = \tilde{f} : \{0,1\}^{O(m)} \to \{0,1\}$. Then, given any size s circuit $\tilde{C}$ such that $\Pr[\tilde{C}(y) = \tilde{f}(y)] \geq 1 - \frac{1}{50}$, we can build a circuit C of size $\leq s \cdot poly(m)$ that exactly computes f.

$\overset{m}{f}$ - worst-case hard. ie Any ckt C makes some mistake $\longrightarrow$ $\tilde{f} = En(f)^{O(m)}$ is $1/50$ - weakly hard.

$\Downarrow$

$O(mk)$ $\tilde{f}^{\oplus k}$ - is strongly hard

$\longleftarrow$ NW

Thm: [Impagliazzo - Wigderson] Suppose $h: \{0,1\}^* \to \{0,1\}$ is a language in $E = TIME(2^{O(n)})$ that <u>cannot</u> be computed by circuits of size $s(n)$.
Then

$$BPP \subseteq TIME\left(2^{O(d(n))} \cdot poly(n)\right)$$

where $d(n) = s^{-1}(poly(n))^2 / \log(n)$.

(same trade-offs like in NW).

$\therefore$ If there are "hard functions", then randomness is "easy"!

Other PRGs from hard functions:

Thm [Umans]: Given a fn $f: \{0,1\}^{\log n} \to \{0,1\}$ with circuit complexity $\geq s$, there is a PRG $G_f: \{0,1\}^{O(\log n)} \to \{0,1\}^m$ against circuits of size $m = s^{\Omega(1)}$.

(optimal hardness-randomness trade offs).

$$G_f(y) = \left(f(y), f(Ay), f(A^2y), \ldots, f(A^{m-1}y)\right)$$

roughly...

Do PRGs imply hardness? Yes! Pset 3.

Next class: What about PRGs against all tests?

ie $$\left| \underset{x \sim \mathcal{U}_m}{E}[A(x)] - \underset{y \sim \mathcal{U}_d}{E}[A(G(y))] \right| \leq \varepsilon$$

for all $A: \{0,1\}^m \to [-1,1]$?

$$\Updownarrow$$

$$G(\mathcal{U}_d) \overset{T.V}{\approx_\varepsilon} \mathcal{U}_m. \quad \longrightarrow \text{Extractors.}$$

extracting randomness from weak sources.