MIT 6.875

Foundations of Cryptography
Lecture 7

Recap + Today

- ✓ Define one-way functions (OWF).
- ✓ Define Hardcore bits (HCB).
- ✓ Show that one-way functions* + HCB $\Rightarrow$ PRG

Goldreich-Levin Theorem: "every OWF has a HCB."

Recap + Today

- ✓ Define one-way functions (OWF).
- ✓ Define Hardcore bits (HCB).
- ✓ Show that one-way functions* + HCB $\Rightarrow$ PRG

Goldreich-Levin Theorem: for every OWF/OWP F, there is another OWF/OWP F' which has a HCB.

Goldreich-Levin (GL) Theorem: Version 1

Let $\{B_r : \{0,1\}^n \rightarrow \{0,1\}\}$ where

$$B_r(x) = \langle r, x \rangle = \sum_{i=1}^{n} r_i x_i \bmod 2$$

be a collection of predicates (one for each r). Then, for ***every*** one-way function F, a ***random*** B_r is hardcore. That is, for every one-way function F, every PPT A, there is a negligible function μ s.t.

$$\Pr[x \leftarrow \{0,1\}^n; r \leftarrow \{0,1\}^n : A(F(x), r) = B_r(x)] \leq \frac{1}{2} + \mu(n)$$

GL Theorem: Version 2

For ***every*** one-way function F, there is a related one-way function

$$F'(x,r) = (F(x),r)$$

which has a *deterministic* hardcore predicate. In particular, the predicate $B(x,r) = \langle r,x\rangle \bmod 2$ is hardcore for F'.

$$\Pr\left[x \leftarrow \{0,1\}^n; r \leftarrow \{0,1\}^n: A\left(F'(x,r)\right) = \langle r,x\rangle\right] \le \frac{1}{2} + \mu(n)$$

Key Point:

This statement is *sufficient* to construct PRGs from any OWP.

If there are OWPs, then there are PRGs

CONSTRUCTION

Let F be a one-way permutation, then G defined below is a PRG.

Then, define $G(x, r) = \mathrm{F}'(\mathrm{x}, \mathrm{r}) \,||\, \langle r, x \rangle = \mathrm{F}(x) \,||\, r \,||\, \langle r, x \rangle$.

Theorem: G is a PRG assuming F is a one-way permutation.

We proved a weaker version in L6:

Let's assume a pretty good predictor P

$$\Pr[x \leftarrow \{0,1\}^n; r \leftarrow \{0,1\}^n : P(F(x), r) = \langle r, x \rangle] \geq \frac{3}{4} + 1/p(n)$$

Then there is a OWF inverter A.

$$\Pr\left[x \leftarrow \{0,1\}^n : A\big(F(x)\big) \in F^{-1}\big(F(x)\big)\right] \geq 1/p'(n)$$

We proved a weaker version in L6:

Let's assume a pretty good predictor P

$$\Pr[x \leftarrow \{0,1\}^n; r \leftarrow \{0,1\}^n : P(F(x), r) = \langle r, x \rangle] \geq \mathbf{\frac{3}{4}} + 1/p(n)$$

First, we used an **averaging argument**.

Claim: For at least a $1/2p(n)$ fraction of the x,

$$\Pr[r \leftarrow \{0,1\}^n : P(F(x), r) = \langle r, x \rangle] \geq \frac{3}{4} + 1/2p(n)$$

Call these **the good x**.

Proof: On the board.

We proved a weaker version in L6:

For at least a $1/2p(n)$ fraction of the x,

$$\Pr[r \leftarrow \{0,1\}^n: P(F(x),r) = \langle r,x\rangle] \geq \frac{3}{4} + 1/2p(n)$$

Key Idea: Linearity

> Pick a random r and ask P to tells us $\langle r,x\rangle$ and $\langle r + e_i, x\rangle$.
> Subtract the two answers to get $\langle e_i, x\rangle = x_i$.

Proof: $\Pr[\text{we compute } x_i \text{ correctly}]$

$$\geq \Pr[\text{P predicts } \langle r,x\rangle \text{ and } \langle r+e_i,x\rangle \text{ correctly}]$$
$$= 1 - \Pr[\text{P predicts} \langle r,x\rangle \text{ or } \langle r+e_i,x\rangle \text{ wrong}]$$
$$\geq 1 - (\Pr[\text{P predicts} \langle r,x\rangle \text{ wrong}] + \Pr[\text{P predicts} \langle r+e_i,x\rangle \text{ wrong}])$$ *(by union bound)*
$$\geq 1 - 2\cdot\left(\frac{1}{4} - \frac{1}{2p(n)}\right) = \frac{1}{2} + 1/p(n)$$

We proved a weaker version in L6:

For at least a $1/2p(n)$ fraction of the x,

$$\Pr[r \leftarrow \{0,1\}^n : P(F(x), r) = \langle r, x \rangle] \geq \frac{3}{4} + 1/2p(n)$$

Inverter A:

Repeat for each $i \in \{1,2,\ldots,n\}$:

Repeat $O(\log n\,(p(n))^2)$ times:

Pick a random r and ask P to tells us $\langle r, x \rangle$ and $\langle r + e_i, x \rangle$.
Subtract the two answers to get a guess for x_i.

Compute the majority of all such guesses and set the bit as x_i

Output the concatenation of all x_i as x.

Analysis: Chernoff + Union Bound

Who's the culprit here?

For at least a $1/2p(n)$ fraction of the x,

$$\Pr[r \leftarrow \{0,1\}^n : P(F(x), r) = \langle r, x \rangle] \geq \frac{3}{4} + 1/2p(n)$$

Pick a random r and ask P to tells us $\langle r, x \rangle$ and $\langle r + e_i, x \rangle$.
Subtract the two answers to get $\langle e_i, x \rangle = x_i$.

Proof: $\Pr[\text{we compute } x_i \text{ correctly}]$

$$\geq \Pr[\text{P predicts } \langle r, x \rangle \text{ and } \langle r + e_i, x \rangle \text{ correctly}]$$

$$= 1 - \Pr[\text{P predicts} \langle r, x \rangle \text{ or } \langle r + e_i, x \rangle \text{ wrong}]$$

$$\geq 1 - (Pr[\text{P predicts} \langle r, x \rangle \text{ wrong}] + Pr[\text{P predicts} \langle r + e_i, x \rangle \text{ wrong}])$$ ***(by union bound)***

$$\geq 1 - 2 \cdot \left(\frac{1}{4} - \frac{1}{2p(n)} \right) = \frac{1}{2} + 1/p(n)$$

The Real Proof of the GL Theorem

(attributed to Charlie Rackoff)

Assume (after averaging) that for $\geq 1/2p(n)$ f

$$\Pr[r \leftarrow \{0,1\}^n : P(F(x), r) = \langle r, x \rangle] \geq \mathbf{\frac{1}{2}} + 1/$$

For a minute, assume we have a bit of help/ad

Pick a random r, ask the Oracle to tells us $\langle r, x \rangle$ and ask P to tell us $\langle r + e_i, x \rangle$. Subtract the two answers to get $\langle e_i, x \rangle = x_i$.

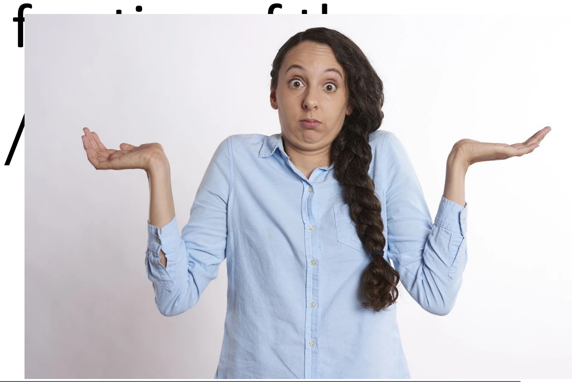
Proof: $\Pr$[we compute x_i correctly]

$\geq$ ***Pr*****[P predicts$\langle r + e_i, x \rangle$ correctly]** $\geq \frac{1}{2} + 1/2p(n)$

The Real Proof of the GL Theorem

Assume (after averaging) that for $\geq 1/2p(n)$ f

$\Pr[r \leftarrow \{0,1\}^n : P(F(x), r) = \langle r, x \rangle] \geq \mathbf{\frac{1}{2}} + 1/$



Pick a random r, **guess** $\langle r, x \rangle$ and ask P to tell us $\langle r + e_i, x \rangle$.
Subtract the two to get $\langle e_i, x \rangle = x_i$.

If our guesses are all correct, then the analysis works out just as before.

But what's the chance…?
The number of r's is $m = O(n \log n\, (p(n))^2)$.

Parsimony in Guessing

Pick random "seed vectors" $s_1, \ldots, s_{\log(m+1)}$, and **guess** $c_j = \langle s_j, x \rangle$ for all j.

The probability that all guesses are correct is $\frac{1}{2^{\log(m+1)}} = 1/(m+1)$ which is not bad.

From the seed vectors, generate many more r_i.

Let $T_1, \ldots, T_m$ denote all possible non-empty subsets of $\{1,2,\ldots,\log(m+1)\}$. We will let

$$r_i = \bigoplus_{j \in T_i} s_j \qquad \text{and} \qquad b_i = \bigoplus_{j \in T_i} c_j$$

Key Observation: If the guesses $c_1, \ldots, c_{\log(m+1)}$ are all correct, then so are the $b_1, \ldots, b_m$.

The OWF Inverter

Generate random $s_1, \dots, s_{\log(m+1)}$ and bits $c_1, \dots, c_{\log(m+1)}$.

From them, derive $r_1, \dots, r_{\log(m+1)}$ and bits $b_1, \dots, b_m$ as in the previous slide.

Repeat for each $i \in \{1,2,\dots,n\}$:

Repeat $\mathbf{100n(p(n))^2}$ times:

Ask P to tells us $\langle r_i + e_i, x \rangle$. XOR P's reply with b_i to get a guess for x_i.

Compute the majority of all such guesses and set the bit as x_i

Output the concatenation of all x_i as x.

Analysis of the Inverter

Let's condition on the guesses $c_1, \ldots, c_{\log(m+1)}$ being all correct.

The main issue: The r_i are not independent (can't do Chernoff)

Key Observation: The r_i **are** pairwise independent.

Therefore, can apply Chebyshev!

We have that
$p :=$ $\Pr[\text{Inverter succeeds} \mid \text{all guesses correct}, \text{good x}] \geq 0.99$.

(Pf. on the board, also in the next two slides)

Putting it all together

Pr[Inverter succeeds]

$\geq$ Pr[Inverter succeeds | all guesses correct, good x] $\cdot$

Pr[all guesses correct] $\cdot$ Pr[good x]

$$= \frac{1}{m+1} \cdot \frac{1}{2p(n)} \cdot p$$

$$= \frac{1}{2n^2 p(n)^3} \cdot p$$

So, it suffices to show that p is large.

By our calculation (on the board), $p \geq 0.99$, so we are done. ■

Can also make the success probability $\approx 1/p(n)$ by enumerating over all the "guesses". Each guess results in a supposed inverse, but we can check which of them is the actual inverse!

The Coding-Theoretic View of GL

$x \to (\langle x, r\rangle)_{r\in\{0,1\}^n}$ can be viewed as a highly redundant, exponentially long encoding of x = **the Hadamard code**.

$P(F(x), r)$ can be thought of as providing access to a **noisy** codeword.

What we proved:

- **unique decoding** algorithm for Hadamard code with error rate $\frac{1}{4} - 1/p(n)$.
- **list-decoding algorithm** for Hadamard code with error rate $\frac{1}{2} - 1/p(n)$.

Hardcore Predicates from any List-Decodable Code

(due to Impagliazzo and Sudan)

$x \rightarrow C(x)$ is the encoding.

Given a $C(x)$ that is incorrect at $\frac{1}{2} - \varepsilon$ fraction of the locations, a list-decoder outputs a list $\{x_1, \ldots, x_m\}$ of possibilities for x.

The hardcore predicate is

$$\boldsymbol{B_i(x) = C(x)_i}.$$

A hardcore-bit predictor gives us access to a corrupted codeword. Running the list-decoder on it gives us the list of possible inverses. The fact that the OWF is easy to compute means that we can filter out the bogus (non-)inverses.

Recap

1. Defined one-way functions (OWF).

2. Defined Hardcore bits (HCB).

3. Goldreich-Levin Theorem: every OWF has a HCB.

   (showed proof for an important special case)

4. Show that one-way *permutations* (OWP) ⇒ PRG

   (in fact, one-way functions ⇒ PRG, but that's a much harder theorem)

Universal Hardcore Predicate Conjecture 1

For every one-way function F,
there exists a circuit B_F s.t.
for every PPT Circuit/Turing Machine A,
there is a negligible function μ s.t.

$$\Pr\left[x \leftarrow \{0,1\}^n : A\big(F(x)\big) = B_F(x)\right] \leq \frac{1}{2} + \mu(n)$$

In fact: I conjecture that for every one-way function F, there ***exists*** an r_F for which the predicate $B_{r_F}(x) = \langle r_F, x\rangle$ that is hardcore.

Universal Hardcore Predicate Conjecture 2

For every one-way function F,
there is **an efficiently generatable** circuit B_F s.t.
for every PPT Circuit/Turing Machine A,
there is a negligible function μ s.t.

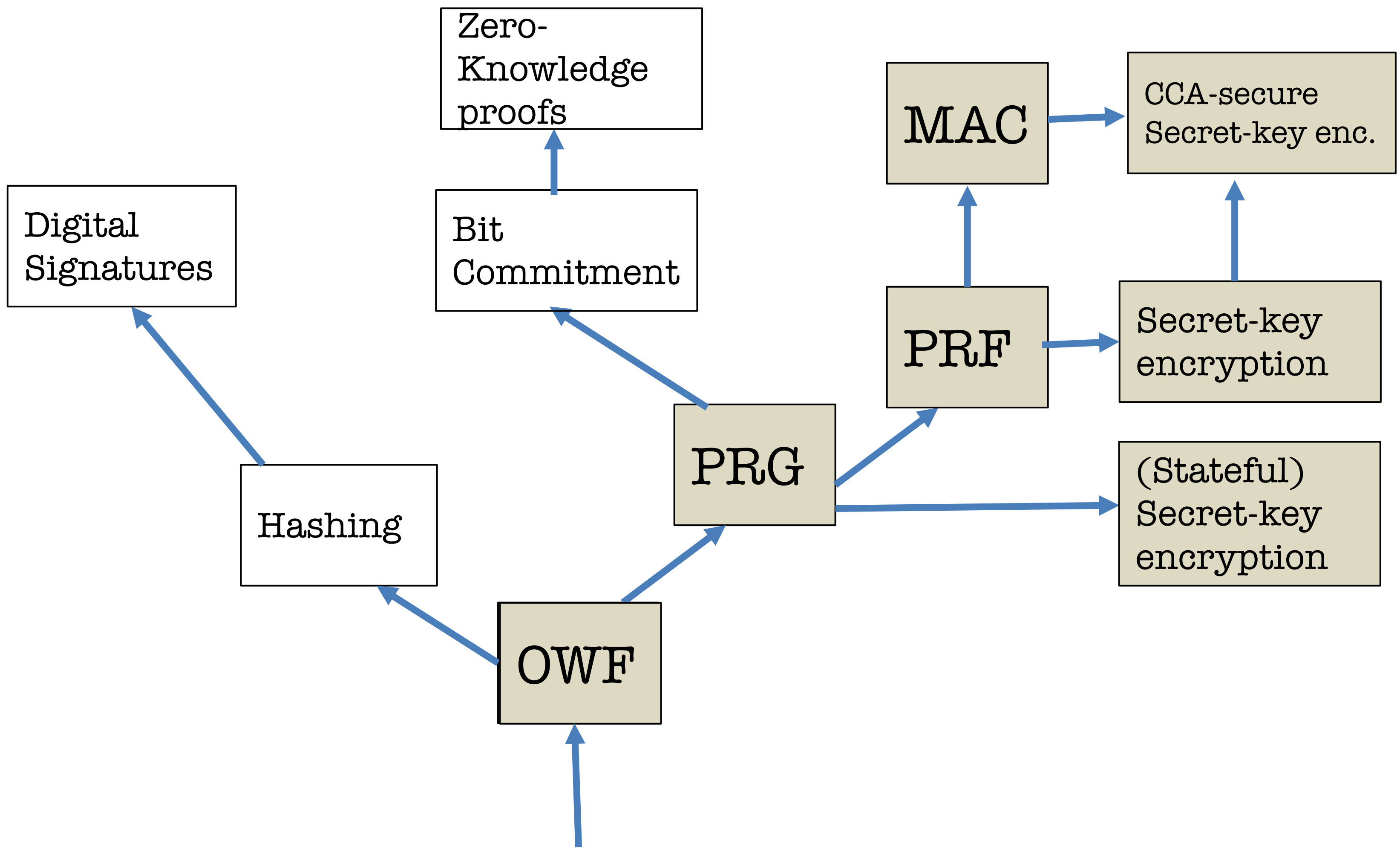
$$\Pr\left[x \leftarrow \{0,1\}^n : A\big(F(x)\big) = B_F(x)\right] \leq \frac{1}{2} + \mu(n)$$

Other Topics (Time permitting)

1. OWF $\Rightarrow$ PRG?

2. Pseudorandom Permutations from Pseudorandom Functions (the Luby-Rackoff construction)

Minicrypt:



Candidate Constructions: from number theory, geometry, combinatorics,...