Recommended reading:

The Goldreich-Levin Theorem

KL, Section 6.3

We have seen that PRGs and permutations with hardcore predicates allow us to construct fascinating cryptographic objects.

We have also seen that PRGs and permutations with hardcore predicates are easy-to-compute but hard-to-invert. → one-way functions (OWFs)

Given that we don't have unconditional constructions of many cryptographic objects, it is natural to try to understand what is the minimal assumption for a class of cryptographic objects.

It turns out that OWFs are a minimal assumption for most of cryptography!

Goldreich-Levin 1989: OWFs ⇒ hardcore predicates

Håstad-Impagliazzo-Levin-Luby 1989/1990: OWFs ⇒ PRGs

This lecture: OWPs $\Rightarrow$ permutations with hardcore predicates

($\Rightarrow$ PRGs)

we don't know how to construct OWPs from OWFs, so this is weaker than results above.

We have some good evidence that constructing an OWP from an arbitrary OWF is impossible.

(see Rudich's 1988 PhD thesis)

Theorem: Suppose that f is an OWP. Then,

$g(x,r) = (f(x),r)$ is an OWP and

$|x| = |r|$

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

is a hardcore predicate for g.

Obs: We will use the fact that $\langle \cdot,\cdot \rangle$ is bilinear, ie,

$$\langle a+b,c \rangle = \langle a,c \rangle + \langle b,c \rangle$$

$$\langle a,b+c \rangle = \langle a,b \rangle + \langle a,c \rangle$$

Proof:

f OWP $\Rightarrow$ g OWP : homework

Suppose that $P(x,r) = \langle x, r \rangle$ is not a hardcore predicate for g. Then, there is a PPT algorithm A and a polynomial $p(n)$ such that for infinitely many n,

$$\Pr_{x,r \leftarrow \{0,1\}^n} \left(A(1^n, f(x), r) = \langle x, r \rangle \right) \geq \frac{1}{2} + \frac{1}{p(n)} . \quad (*)$$

We will use A to design another PPT algo A' that inverts f with non-negligible probability.

First, note that $(*)$ only provides guarantees for a unif. random x. But, with some hindsight, we will make several queries to A on the same x, and so would like a guarantee that holds for fixed x.

Towards this, define

$$GOOD = \left\{ x : \Pr_{r \leftarrow \{0,1\}^n} \left(A(1^n, f(x), r) = \langle x, r \rangle \right) \geq \frac{1}{2} + \frac{1}{2p(n)} \right\},$$

the set of "good" x's on which A succeeds with decent advantage.

Claim: $|GOOD| \geq \frac{2^n}{2p(n)}$.

In other words, a non-negligible fraction of x's are good.

Proof: This follows by an averaging argument.

Suppose that $|GOOD| < \frac{2^n}{2p(n)}$. Then,

$$\Pr_{x,r \leftarrow \{0,1\}^n}\left(A(1^n, f(x), r) = \langle x, r\rangle\right) \leq \Pr_{x \leftarrow \{0,1\}^n}(x \in GOOD) + \left(\frac{1}{2} + \frac{1}{2p(n)}\right)$$

$$< \frac{1}{2p(n)} + \frac{1}{2} + \frac{1}{2p(n)}$$

$$= \frac{1}{2} + \frac{1}{p(n)}. \quad \leftarrow \text{Contradiction to } (*)$$

$\blacksquare$

Warmup 1: $\Pr_{r \leftarrow \{0,1\}^n}\left(A(1^n, f(x), r) = \langle x, r\rangle\right) = 1$ for $x \in GOOD$

In this case we can efficiently find x given $f(x)$, if $x \in GOOD$, bit by bit:

- Let $e_i = (0, 0, \ldots, 0, \underset{\text{i-th coordinate}}{1}, 0, \ldots, 0)$
- $A(1^n, f(x), e_i) = \langle x, e_i\rangle = x_i$
- Repeating this for $i = 1, \ldots, n$ recovers x.

$\Rightarrow$ invert successfully with probability $\geq \frac{|GOOD|}{2^n} \geq \frac{1}{2p(n)}$ $\blacksquare$

Warmup 2: $\Pr_{r \leftarrow \{0,1\}^n}\left(A(1^n, f(x), r) = \langle x, r\rangle\right) \geq 3/4 + \frac{1}{p(n)}$ $\forall x \in \text{GOOD}$

Key observation: If we know $\langle x, r\rangle$ and $\langle x, r+e_i\rangle$, then we know $\langle x, r\rangle + \langle x, r+e_i\rangle = \langle x, r+r+e_i\rangle = \langle x, e_i\rangle = x_i$.

- First, let's show that we can guess x_i with probability at least $\frac{1}{2} + \frac{2}{p(n)}$.

  → Sample $r_i \leftarrow \{0,1\}^n$

  → Compute $b_i' \leftarrow A(1^n, f(x), r^{(i)})$

  $b_i'' \leftarrow A(1^n, f(x), r^{(i)} + e_i)$

  → Output $b_i = b_i' + b_i''$.

$$\Pr(b_i \neq x_i) \leq \Pr\left(b_i' \neq \langle x, r^{(i)}\rangle \text{ or } b_i'' \neq \langle x, r^{(i)} + e_i\rangle\right)$$

union bound $\leq \Pr\left(b_i' \neq \langle x, r^{(i)}\rangle\right) + \Pr\left(b_i'' \neq \langle x, r^{(i)} + e_i\rangle\right)$

$\leq \frac{1}{4} - \frac{1}{p(n)}$ $\quad$ $\leq \frac{1}{4} - \frac{1}{p(n)}$

because $r^{(i)}$ is unif. random and so is $r^{(i)} + e_i$

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

- Now, let's amplify the probability of successfully guessing x_i. We will use a Chernoff bound.

Repeat the procedure above $s(n) = n\,p(n)^2$ times, each time with a freshly sampled r_i, then take majority.

Let $b_i^{(j)}$ be the guess for x_i in the j-th trial.

Define $Z_i^{(j)} = 1_{\{b_i^{(j)} = x_i\}}$.

Then, the $Z_i^{(j)}$'s are independent (because x is fixed and the $r_i^{(j)}$'s are independent), and

$$\Pr(Z_i^{(j)} = 1) \geq \frac{1}{2} + \frac{2}{p(n)}.$$

Define $Z_i = \sum_{j=1}^{s(n)} Z_i^{(j)}$. We guess x_i if $Z_i > \frac{s(n)}{2}$.

We have $E(Z_i) \geq s(n) \cdot \left(\frac{1}{2} + \frac{2}{p(n)}\right)$, so

$$\Pr\left(Z_i \leq \frac{s(n)}{2}\right) \leq \Pr\left(Z_i \leq E(Z_i) - \frac{2s(n)}{p(n)}\right)$$

Chernoff bound $\longleftarrow \leq 2e^{-s(n)/p(n)^2}$

$$= 2e^{-n}.$$

- We saw above that for each i we can efficiently produce a guess b_i of x_i that is correct with prob $\geq 1 - 2e^{-n}$.

To invert x we must produce correct guesses for $x_1, \ldots, x_n$.

By a union bound,

$$\begin{aligned} \Pr(b_i \neq x_i \text{ for some } i) &\leq \sum_{i=1}^{n} \Pr(b_i \neq x_i) \\ &\leq \sum_{i=1}^{n} 2e^{-n} \\ &= 2ne^{-n}, \end{aligned}$$

which is negligible. So we invert $f(x)$ with probability at least

$$|GOOD|/2^n \cdot (1 - 2ne^{-n}) \to \text{non-negligible.}$$ ▨

Full proof: $\Pr_{r \leftarrow \{0,1\}^n}\left(A(1^n, f(x), r) = \langle x, r\rangle\right) \geq \frac{1}{2} + \frac{1}{p(n)} \quad \forall x \in GOOD$

In the previous case we produced a guess for x_i by querying A on r_i and $r_i + e_i$, for random r_i.

However, this doesn't work now. For example, A may always guess correctly if $(r_i)_i = 0$ and guess incorrectly with prob. $1 - 1/n$ if $(r_i)_i = 1$.
This means that our guess would be wrong with prob $1 - 1/n$, while a query on a random r is still correct with probability $1/2 + \frac{1}{2n}$.

The issue that causes this is that the queries r_i and $r_i + e_i$ are not independent.

How can we make these queries independent?

Here's an approach that doesn't work:

- Sample $r_i \leftarrow \{0,1\}^n$.
- Sample $b' \leftarrow \{0,1\}$ as a guess for $\langle x, r\rangle$, correct with prob. $1/2$.
- Query $b'' \leftarrow A(1^n, f(x), r_i + e_i)$, guess of $\langle x, r + e_i\rangle$ correct with prob $1/2 + \frac{1}{p(n)}$.

If we follow the strategy of Warmup 2 then we guess x_i with high prob, provided that our guesses for $\langle x, r_i^{(j)} \rangle$ are correct for all $j = 1, \ldots, s(n)$.
The probability that this happens is $2^{-s(n)}$, which is negligible...

How can we overcome this barrier?

Key observation: Suppose that we have correctly guessed $c_1 = \langle x, r_1 \rangle, \ldots, c_\ell = \langle x, r_\ell \rangle$ for some $r_1, \ldots, r_\ell$.

Then,

$$c_I = \sum_{i \in I} c_i = \langle x, \underbrace{\sum_{i \in I} r_i}_{r_I} \rangle$$

for all sets $I \subseteq \{1, \ldots, \ell\}$, $I \neq \emptyset$.

This means that we can turn ℓ correct guesses into $2^\ell - 1$ correct guesses!

BUT, the r_I's are not independent... we needed this independence for the strategy in Warmup 2.

Key observation: If $r_1, \ldots, r_\ell$ are independent and uniformly random, then the r_I's are pairwise independent.

So here's the final strategy to guess x_i with high prob:

→ Set $s(n) = n \cdot p(n)^2$, $\ell = \lceil \log(s(n)+1) \rceil$

→ Sample $r_1, \ldots, r_\ell \leftarrow \{0,1\}^n$ indep.

→ Sample $c_1, \ldots, c_\ell \leftarrow \{0,1\}$ indep → these are guesses for $\langle x, r_i \rangle$.

→ For each $I \subseteq \{1, \ldots, \ell\}$, $I \neq \emptyset$:

→ Compute $b'_I \leftarrow A(1^n, f(x), r_I + c_i)$

→ Set $b_I = c_I + b'_I$.

→ Output guess $x'_i = \text{majority}(b_I)_{\substack{I \subseteq \{1,\ldots,\ell\} \\ I \neq \emptyset}}$

Fix $x \in GOOD$. We now upper bound $\Pr[x'_i \neq x_i \mid c_j = \langle x, r_j \rangle \; \forall j \in \{1,\ldots,\ell\}]$

Let $Z_I = 1\{b_I = x_i\}$ and $Z = \sum_{\substack{I \subseteq \{1,\ldots,\ell\} \\ I \neq \emptyset}} Z_I$.

Then we must upper bound $\Pr\left(Z \leq \frac{2^\ell - 1}{2}\right)$.

In Warmup 2 we used a Chernoff bound to do this. But here the r_I's are not independent. However, they are pairwise independent, so we can use Chebyshev's inequality.

Lemma: If $Z = \sum_{j=1}^{s} z_j$ with the z_j's pairwise independent and $Pr(z_j = 1) = \sigma$ for all j, then

$$Pr\left(|Z - E(Z)| \geq \delta \cdot s \right) \leq \frac{1}{4\delta^2 s}.$$

Proof: Homework. Apply Chebyshev's inequality.

Since in our case $E(Z) = \sum_I E(z_I) = (2^{\ell-1})\sigma = s(n)$

for $\sigma \geq 1/2 + 1/p(n)$ and the z_I's are pairwise independent, we can use the lemma to conclude that

$$Pr\left(Z \leq \frac{2^{\ell-1}}{2} \right) \leq Pr\left(|Z - E(Z)| \leq \frac{2^{\ell-1}}{p(n)} \right)$$

$$\leq \frac{p(n)^2}{4(2^{\ell-1})} \leq \frac{1}{4n}.$$

Here, we used our choice $s(n) = n \cdot p(n)^2$.

This means that we recover x_i except with probability $\leq 1/4n$.

Now, repeat this procedure with the <u>same</u> r_j's and c_j's for all $i = 1, \ldots, n$.

By a union bound, the prob. that it fails for some i is at most $n \cdot \frac{1}{4n} = 1/4$.

Therefore, we successfully recover x with prob. at least

$$\Pr(x \in GOOD) \cdot \Pr(c_j = \langle x, r_j \rangle \; \forall j \in \{1, \ldots, \ell\}) \cdot 3/4$$

$$\geq \frac{1}{2p(n)} \cdot 2^{-\ell} \cdot 3/4$$

$$\geq \frac{1}{2p(n)} \cdot \frac{1}{2 \cdot s(n)} \cdot 3/4$$

$$= \frac{3}{16 \, n \, p(n)^3},$$

which is non-negligible. ▨