One Way Functions (Part II)

601.642/442: Modern Cryptography

Fall 2018

Last Time

- Modeling adversaries as non-uniform PPT Turing machines
- Negligible and noticeable functions
- Definitions of strong and weak OWFs
- Factoring assumption
- Candidate weak OWF $f_\times$ based on factoring assumption

Today's Agenda

- Proving $f_\times$ is a weak OWF
- Yao's hardness amplification: from weak to strong OWFs

Recall

Definition (Weak One Way Function)

A function $f : \{0,1\}^* \to \{0,1\}^*$ is a *weak one-way function* if it satisfies the following two conditions:

- **Easy to compute:** there is a PPT algorithm $\mathcal{C}$ s.t. $\forall x \in \{0,1\}^*$,

$$\Pr\left[\mathcal{C}(x) = f(x)\right] = 1.$$

- **Somewhat hard to invert:** there is a noticeable function $\varepsilon : \mathbb{N} \to \mathbb{R}$ s.t. for every non-uniform PPT $\mathcal{A}$ and $\forall n \in \mathbb{N}$:

$$\Pr\left[x \leftarrow \{0,1\}^n, x' \leftarrow \mathcal{A}(1^n, f(x)) : f(x') \neq f(x)\right] \geqslant \varepsilon(n).$$

Noticeable (or non-negligible): $\exists c$ s.t. for infinitely many $n \in \mathbb{N}$, $\varepsilon(n) \geqslant \frac{1}{n^c}$.

Recall (contd.)

- Multiplication function $f_\times : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$:

$$f_\times(x, y) = \begin{cases} \perp & \text{if } x = 1 \vee y = 1 \\ x \cdot y & \text{otherwise} \end{cases}$$

Theorem

Assuming the factoring assumption, function $f_\times$ is a weak OWF.

Proof Idea

- Let $\textsc{Good}$ be the set of inputs (x, y) to $f_\times$ s.t. both x and y are prime numbers
- When $(x, y) \in \textsc{Good}$, adversary cannot invert $f_\times(x, y)$ (due to hardness of factoring)
- Suppose adversary inverts with probability 1 when $(x, y) \notin \textsc{Good}$
- But if $\Pr[(x, y) \in \textsc{Good}]$ is noticeable, then overall, adversary can only invert with a bounded noticeable probability
- Formally: Let $q(n) = 8n^2$. Will show that no non-uniform PPT adversary can invert $f_\times$ with probability greater than $1 - \frac{1}{q(n)}$

Proof via Reduction

Goal: Given an adversary A that breaks weak one-wayness of $f_\times$ with probability *at least* $1 - \frac{1}{q(n)}$, we will construct an adversary B that breaks the factoring assumption with non-negligible probability

Adversary $B(z)$**:**

1. $x, y \overset{\$}{\leftarrow} 0, 1^n$
2. If x and y are primes, then $z' = z$
3. Else, $z' = x \cdot y$
4. $w \leftarrow A(1^n, z')$
5. Output w if x and y are primes

Analysis of B:

- Since A is non-uniform PPT, so is B (using polynomial-time primality testing)
- A fails to invert with probability at most $\frac{1}{q(n)} = \frac{1}{8n^2}$
- B fails to pass z to A with probability at most $1 - \frac{1}{4n^2}$ (by Chebyshev's Thm.)
- Union bound: B fails with probability at most $1 - \frac{1}{8n^2}$
- B succeeds with probability at least $\frac{1}{8n^2}$: Contradiction to factoring assumption!

Weak to Strong OWFs

Theorem (Yao)

Strong OWFs exist if and only weak OWFs exist

- This is called hardness amplification: convert a somewhat hard problem into a really hard problem

Weak to Strong OWFs

Theorem (Yao)

Strong OWFs exist if and only weak OWFs exist

- This is called hardness amplification: convert a somewhat hard problem into a really hard problem
- Intuition: Use the weak OWF *many* times
- Think: Is $f(f(\ldots f(x)))$ a good idea?

Weak to Strong OWFs

Theorem

For any weak one-way function $f : \{0,1\}^n \to \{0,1\}^n$, there exists a polynomial $N(\cdot)$ s.t. the function $F : \{0,1\}^{n \cdot N(n)} \to \{0,1\}^{n \cdot N(n)}$ defined as

$$F(x_1, \ldots, x_N(n)) = (f(x_1), \ldots, f(x_N(n)))$$

is strongly one-way.

- <u>Think:</u> Show that when f is the $f_\times$ function, then F is a strong one-way function

Weak to Strong OWFs: Intuition

- Good inputs: hard to invert, Bad inputs: easy to invert
- A OWF is weak when the fraction of Bad inputs is **noticeable**
- In a strong OWF, the fraction of Bad inputs is **negligible**
- To convert weak OWF to strong, use the weak OWF on **many** (say N) inputs independently
- In order to successfully invert the new OWF, adversary must invert ALL the N outputs of the weak OWF
- If N is sufficiently large and the inputs are chosen independently at random, then the probability of inverting all of them should be small

Weak to Strong OWFs: Intuition

- The above intuition does not quite work as you expect because even though the instances are chosen independently, adversary gets to see them all together and does not have to invert them independently.
- Nevertheless, it can be shown via a non-trivial proof that hardness does amplify for one-way functions (albeit not all the way to exponentially small inversion probability – there are counterexamples to this!)
- In fact, hardness amplification is not a general phenomenon; for other cases such as interactive arguments (we will study later), hardness does not amplify in general

Weak to Strong OWFs: Proof Strategy

- Set $N = 2nq(n)$
- Since f is weakly one-way, let $q(\cdot)$ be a polynomial s.t. for any PPT adversary, the probability of inverting f is at most $1 - \frac{1}{q(n)}$
- Suppose F is not a strong OWF. Then there exists a PPT adversary A and polynomial $p'(\cdot)$ s.t. A inverts F with probability at least $\frac{1}{p'(nN)} = \frac{1}{p(n)}$
- Goal: Use A to build a PPT adversary B that succeeds in inverting f with probability $> 1 - \frac{1}{q(n)}$ (to derive contradiction)
- Think: How to use A to construct B?
  - Feed input $(y, \ldots, y)$ to A?
  - Feed input $(y, y_2, \ldots, y_N)$ to A where $y_2, \ldots, y_N$ are computed using randomly chosen $x_2, \ldots, x_N$?

Adversary B for f

Adversary $B_i(f, y)$**:**

- Let $y_i = y$
- For every $j \neq i$, sample $x_j \in \{0,1\}^N$ and let $y_j = f(x_j)$
- Let $(z_1, \ldots, z_N) \leftarrow A(1^{nN}, y_1, \ldots, y_N)$
- If $f(z_i) = y$, output z_i, else output $\perp$

Adversary $B(y)$**:**

- For every $i \in N$, run $B_i(f, y)$ $2nNp(n)$ times and output the first non-$\perp$ answer

Analysis of B

Strategy:

- Define $\textsc{Good}_i$ as the set of inputs x to f s.t. B_i inverts $f(x)$ with noticeable probability $\geqslant \frac{1}{2Np(n)}$ (where probability is over randomness of B_i)
- Claim 1: There exists $i \in N$ s.t. $x \in \textsc{Good}_i$ with probability at least $1 - \frac{1}{2q(n)}$
- Claim 2: When $x \in \textsc{Good}_i$ (for any $i \in N$), B fails to invert $f(x)$ with negligible probability
- Claim 3: B's failure probability is at most $\frac{1}{q(n)}$. This means that B succeeds with probability $> 1 - \frac{1}{q(n)}$

Think: Details?