CS 276
Lecture Notes On Graduate Cryptography

This draft is continually being updated.

Sanjam Garg et al.[1]
University of California, Berkeley

Fall 2024

[1] These lecture notes are based on scribe notes taken by students in CS 276 over the years. Also, thanks to Peihan Miao, Akshayaram Srinivasan, and Bhaskar Roberts for helping to improve these notes.

Contents

Chapter 1

One-Way Functions

Information-theoretic cryptography, however desirable, is often not achievable or unsuitable for practical scenarios. For example, one-time pad — an information-theoretically secure cipher — is not very useful because it requires very large keys. Thus, in modern cryptography, we often need to limit attackers' ability. Specifically, (i) we limit security to settings where the attacker can only run in a "reasonable" amount of time, and (ii) the scheme is allowed to fail with some "very small" probability. Furthermore, to leverage the reduced adversarial capabilities, we must make certain computational intractability assumptions.

The existence of one-way functions is one of the weakest computational hardness assumptions (and also the simplest) of interest in cryptography. In this chapter, we study these objects. To define one-way functions, we need new formal vocabulary that we describe in Sections 1.1 and 1.2. Finally, we define one-way functions in Section 1.3.

1.1 Probabilistic Polynomial Time

A polynomial time Turing Machine is one which halts in time polynomial in its input length. A probabilistic Turing Machine is allowed to make random choices in its execution. More formally:

Definition 1.1 (Probabilistic Polynomial Time (PPT)) *A Turing Machine M is said to be a PPT Turing Machine if $\exists c \in \mathbb{Z}^+$ such that $\forall x \in \{0,1\}^*$, $M(x)$ halts in $|x|^c$ steps.*

A *non-uniform* PPT Turing Machine is a collection of machines one for each input length, as opposed to a single machine that must work for all input lengths.

Definition 1.2 (Non-uniform PPT) *A non-uniform PPT machine is a sequence of Turing Machines $\{M_1, M_2, \cdots\}$ such that $\exists c \in \mathbb{Z}^+$ such that $\forall x \in \{0,1\}^*$, $M_{|x|}(x)$ halts in $|x|^c$ steps.*

1.2 Noticeable and Negligible Functions

Noticeable and negligible functions are used to characterize the "largeness" or "smallness" of a function describing the probability of some event. Intuitively, a noticeable function is required to be larger than some inverse-polynomially function in the input parameter. On the other hand, a negligible function must be smaller than any inverse-polynomial function of the input parameter. More formally:

Definition 1.3 (Noticeable Function) *A function $\mu(\cdot) : \mathbb{Z}^+ \to [0,1]$ is noticeable iff $\exists c, n_0 \in \mathbb{Z}^+$ such that $\forall n \geq n_0,\ \mu(n) \geq n^{-c}$.*

Example. Observe that $\mu(n) = n^{-3}$ is a noticeable function. (Notice that the above definition is satisfied for $c = 3$ and $n_0 = 1$.)

Definition 1.4 (Negligible Function) *A function $\mu(\cdot) : \mathbb{Z}^+ \to [0,1]$ is negligible iff $\forall c \in \mathbb{Z}^+\ \exists n_0 \in \mathbb{Z}^+$ such that $\forall n \geq n_0,\ \mu(n) < n^{-c}$.*

Example. $\mu(n) = 2^{-n}$ is an example of a negligible function. This can be observed as follows. Consider an arbitrary $c \in \mathbb{Z}^+$ and set $n_0 = c^2$. Now, observe that for all $n \geq n_0$, we have that $\frac{n}{\log_2 n} \geq \frac{n_0}{\log_2 n_0} \geq \frac{n_0}{\sqrt{n_0}} = \sqrt{n_0} = c$. This allows us to conclude that

$$\mu(n) = 2^{-n} = n^{-\frac{n}{\log_2 n}} \leq n^{-c}.$$

Thus, we have proved that for any $c \in \mathbb{Z}^+$, there exists $n_0 \in \mathbb{Z}^+$ such that for any $n \geq n_0$, $\mu(n) \leq n^{-c}$.

Remark 1.1 *Gap between Noticeable and Negligible Functions. At first thought it might seem that a function that is not negligible (or, a non-negligible function) must be a noticeable. This is not true! Negating the definition of a negligible function, we obtain that a non-negligible function $\mu(\cdot)$ is such that $\exists c \in \mathbb{Z}^+$ such that $\forall n_0 \in \mathbb{Z}^+$, $\exists n \geq n_0$ such that $\mu(n) \geq n^{-c}$. Note that this requirement is satisfied as long as $\mu(n) \geq n^{-c}$ for infinitely many choices of $n \in \mathbb{Z}^+$. However, a noticeable function requires this condition to be true for every $n \geq n_0$.*

Below we give example of a function $\mu(\cdot)$ that is neither negligible nor noticeable.

$$\mu(n) = \begin{cases} 2^{-n} & : x \mod 2 = 0 \\ n^{-3} & : x \mod 2 \neq 0 \end{cases}$$

This function is obtained by interleaving negligible and noticeable functions. It cannot be negligible (resp., noticeable) because it is greater (resp., less) than an inverse-polynomially function for infinitely many input choices.

Properties of Negligible Functions. Sum and product of two negligible functions is still a negligible function. We argue this for the sum function below and defer the problem for products to Exercise 1.2.

Lemma 1.1 *If $\mu(n)$ and $\nu(n)$ are negligible functions then*

1. *$\psi(n) = \min\{\mu(n) + \nu(n), 1\}$ is a negligible function,*
2. *$\psi(n) = \mu(n) \cdot \nu(n)$ is a negligible function, and*
3. *$\psi(n) = \mathsf{poly}(\mu(n))$ is[1] a negligible function.*

[1]Where $\mathsf{ploy}(n)$ is an unspecified polynomial function.

Proof. of (i) We need to show that for any $c \in \mathbb{Z}^+$, we can find n_0 such that $\forall n \geq n_0$, $\psi(n) \leq n^{-c}$. Our argument proceeds as follows. Given the fact that μ and ν are negligible we can conclude that there exist n_1 and n_2 such that $\forall n \geq n_1$, $\mu(n) \leq n^{-(c+1)}$ and $\forall n \geq n_2$, $g(n) \leq n^{-(c+1)}$. Combining the above two fact and setting $n_0 = \max(n_1, n_2, 2)$ we have that for every $n \geq n_0$, $\psi(n) = \mu(n) + \nu(n) \leq n^{-(c+1)} + n^{-(c+1)} = 2n^{-(c+1)} \leq n \cdot n^{-(c+1)}$ (since, $2 \leq n_0 \leq n$). Thus, $\psi(n) \leq n^{-c}$ and hence is negligible. ■

1.3 One-Way Functions

A one-way function $f : \{0,1\}^n \to \{0,1\}^m$ is a function that is easy to compute (computable by a polynomial time machine) but hard to invert. We formalize this by saying that there *cannot* exist a machine that can invert f in polynomial time.

Definition 1.5 (One-Way Functions) *A function $f : \{0,1\}^* \to \{0,1\}^*$ is said to be one-way function if:*

- ***Easy to Compute:*** *$\exists$ a (deterministic) polynomial time machine M such that $\forall x \in \{0,1\}^*$ we have that*

$$M(x) = f(x)$$

- ***Hard to Invert:*** *$\forall$ non-uniform PPT adversary $\mathcal{A}$ we have that*

$$\mu_{\mathcal{A},f}(n) = \Pr_{x \xleftarrow{\$} \{0,1\}^n} [f(\mathcal{A}(1^n, f(x))) = f(x)] \tag{1.1}$$

is a negligible function, $x \xleftarrow{\$} \{0,1\}^n$ denotes that x is drawn uniformly at random from the set $\{0,1\}^n$ and the probability if over the random choices of x and the random coins of $\mathcal{A}$.

The above definition is rather delicate. We next describe problems in the slight variants of this definition that are insecure.

1. What if we require that $\Pr_{x \xleftarrow{\$} \{0,1\}^n}[\mathcal{A}(1^n, f(x)) \in f^{-1}(f(x))] = 0$ instead of being negligible?

   This condition is false for every function f. An adversary $\mathcal{A}$ that outputs an arbitrarily fixed value x_0 succeed in outputting a value in $f^{-1}(f(x))$ with probability at least $1/2^n$.

2. What if we drop the input 1^n to $\mathcal{A}$ in Equation 1.1?

   Consider the function $f(x) = |x|$. In this case, we have that $m = \log_2 n$, or $n = 2^m$. Intuitively, f should not be considered a one-way function, because it is easy to invert f. Namely, given a value y any x such that $|x| = y$ is such that $x \in f^{-1}(y)$. However, according to this definition the adversary gets an m bit string as input, and hence is restricted to running in time polynomial in m. Since each possible x is of size $n = 2^m$, the adversary doesn't even have enough time to write down the answer! Thus, according to the flawed definition above, f would be a one-way function.

   Providing the attacker with 1^n (n repetitions of the 1 bit) as additional input avoids this issue. In particular, it allows the attacker to run in time polynomial in m and n.

A Candidate One-way Function. It is not known whether one-way functions exist. The existence of one-way functions would imply that $P \neq NP$ (see Exercise 1.3), and so of course we do not know of any concrete functions that have been proved to be one-way.

However, there are candidates of functions that could be one-way functions. One example is based on the hardness of factoring. Multiplication can be done easily in $O(n^2)$ time, but so far no polynomial time algorithm is known for factoring.

One candidate might be to say that given an input x, split x into its left and right halves x_1 and x_2, and then output $x_1 \times x_2$. However, this is not a one-way function, because with probability $\frac{3}{4}$, 2 will be a factor of $x_1 \times x_2$, and in general the factors are small often enough that a non-negligible number of the outputs could be factored in polynomial time.

To improve this, we again split x into x_1 and x_2, and use x_1 and x_2 as seeds in order to generate large primes p and q, and then output pq. Since p and q are primes, it is hard to factor pq, and so it is hard to retrieve x_1 and x_2. This function is believed to be one-way.

1.4 Composability of One-Way Functions

Given a one-way function $f : \{0,1\}^n \to \{0,1\}^n$, is the function $f^2(x) = f(f(x))$ also a one-way function? Intuitively, it seems that if it is hard to invert $f(x)$, then it would be just as hard to invert $f(f(x))$. However, this intuition is incorrect and highlights the delicacy when working with cryptographic assumptions and primitives. In particular, assuming one-way functions exists we describe a one-way function $f : \{0,1\}^{n/2} \times \{0,1\}^{n/2} \to \{0,1\}^n$ such that f^2 can be efficiently inverted. Let $g : \{0,1\}^n \to \{0,1\}^n$ be a one-way function then we set f as follows:

$$f(x_1, x_2) = \begin{cases} 0^n & : \text{if } x_1 = 0^{n/2} \\ 0^{n/2} \| g(x_2) & : \text{otherwise} \end{cases}$$

Two observations follow:

1. f^2 is not one-way. This follows from the fact that for all inputs x_1, x_2 we have that $f^2(x_1, x_2) = 0^n$. This function is clearly not one-way!

2. f is one-way. This can be argued as follows. Assume that there exists an adversary $\mathcal{A}$ such that $\mu_{\mathcal{A},f}(n) = \Pr_{x \xleftarrow{\$} \{0,1\}^n}[\mathcal{A}(1^n, f(x)) \in f^{-1}(f(x))]$ is non-negligible. Using such an $\mathcal{A}$ we will describe a construction of adversary $\mathcal{B}$ such that $\mu_{\mathcal{B},g}(n) = \Pr_{x \xleftarrow{\$} \{0,1\}^n}[\mathcal{B}(1^n, g(x)) \in g^{-1}(g(x))]$ is also non-negligible. This would be a contradiction thus proving our claim.

   Description of $\mathcal{B}$: $\mathcal{B}$ on input $y \in \{0,1\}^n$ outputs the n lower-order bits of $\mathcal{A}(1^{2n}, 0^n \| y)$.

   Observe that if $\mathcal{A}$ successfully inverts f then we have that $\mathcal{B}$ successfully inverts g. More formally, we have that:

$$\mu_{\mathcal{B},g}(n) = \Pr_{x \xleftarrow{\$} \{0,1\}^n}\left[\mathcal{A}(1^{2n}, 0^n || g(x)) \in \{0,1\}^n || g^{-1}(g(x))\right].$$

Note that

$$\begin{aligned}
\mu_{\mathcal{A},f}(2n) = & \Pr_{x_1,x_2 \xleftarrow{\$} \{0,1\}^{2n}} [\mathcal{A}(1^{2n}, f(x_1,x_2)) \in f^{-1}(f(\tilde{x}))] \\
\leq & \Pr_{x_1 \xleftarrow{\$} \{0,1\}^n} [x_1 = 0^n] + \Pr_{x_1 \xleftarrow{\$} \{0,1\}^n} [x_1 \neq 0^n] \Pr_{x_2 \xleftarrow{\$} \{0,1\}^n} [\mathcal{A}(1^{2n}, 0^n || g(x_2)) \in \{0,1\}^n || g^{-1}(g(x_2))] \\
= & \frac{1}{2^n} + \left(1 - \frac{1}{2^n}\right) \cdot \Pr_{x_2 \xleftarrow{\$} \{0,1\}^n} [\mathcal{A}(1^{2n}, 0^n || g(x_2)) \in \{0,1\}^n || g^{-1}(g(x_2))] \\
= & \frac{1}{2^n} + \left(1 - \frac{1}{2^n}\right) \cdot \mu_{\mathcal{B},g}(n).
\end{aligned}$$

Rewriting the above expression, we have that $\mu_{\mathcal{B},g}(n) = \frac{\mu_{\mathcal{A},f}(2n) - \frac{1}{2^n}}{1 - \frac{1}{2^n}}$ which is non-negligible as long as $\mu_{\mathcal{A},f}(2n)$ is non-negligible.

1.5 Examples

The goal of this section is to illustrate the general strategy for the problems of the form,

"If f is one-way function, then show that f' (derived from f) is not a one-way function"

Some of the examples include:

- If f is a one-way function, prove that f' defined as $f(f(\cdot))$ is not one-way.
- If f is a one-way function, prove that f' defined by dropping the first bit the output of f is not one-way.

In order to give such a proof, we need to give an example of an one-way function f and show that f' (derived from f) is not one-way. The general strategy for these types of problems is the following:

1. Come up with a contrived function g and show that g is one-way.
2. Construct the new function g' that is derived from g.
3. Show that g' can be inverted with non-negligible probability and thus show that g' is not one-way.

The reason why we need to come-up with a contrived function is that for specific one-way function f, f' (derived from f) could be one-way. To see why this is the case, consider a one-way function $f : \{0,1\}^n \to \{0,1\}^n$ that is additionally injective. Then, one can show that $f^2(\cdot)$ is in fact a one-way function.[2] On the other hand, in the previous section, we showed that there exists a (contrived) function g such that g is one-way but g^2 is not one-way. Hence, we might not always be able to start from any one-way function f and show that f' (derived from f) is not one-way. The first step where we come up a suitable g requires some ingenuity. Once that is done, the second and the third steps would generally be not so hard.

To illustrate these three steps, let us consider a concrete example. We want to show that if f is one-way then f' that is defined by dropping the first bit of the output of f is not one-way.

[2] Try to prove this!

Step-1: Designing the function g. We want to come up with a (contrived) function g and prove that it is one-way. Let us assume that there exists a one-way function $h : \{0,1\}^n \to \{0,1\}^n$. We define the function $g : \{0,1\}^{2n} \to \{0,1\}^{2n}$ as follows:

$$g(x\|y) = \begin{cases} 0^n\|y & \text{if } x = 0^n \\ 1\|0^{n-1}\|g(y) & \text{otherwise} \end{cases}$$

Claim 1.1 *If h is a one-way function, then so is g.*

Proof. Assume for the sake of contradiction that g is not one-way. Then there exists a polynomial time adversary $\mathcal{A}$ and a non-negligible function $\mu(\cdot)$ such that:

$$\Pr_{x,y}[\mathcal{A}(1^n, g(x\|y)) \in g^{-1}(g(x\|y))] = \mu(n)$$

We will use such an adversary $\mathcal{A}$ to invert h with some non-negligible probability. This contradicts the one-wayness of h and thus our assumption that g is not one-way function is false.

Let us now construct an $\mathcal{B}$ that uses $\mathcal{A}$ and inverts h. $\mathcal{B}$ is given $1^n, h(y)$ for a randomly chosen y and its goal is to output $y' \in h^{-1}(h(y))$ with some non-negligible probability. $\mathcal{B}$ works as follows:

1. It samples $x \leftarrow \{0,1\}^n$ randomly.
2. If $x = 0^n$, it samples a random $y' \leftarrow \{0,1\}^n$ and outputs it.
3. Otherwise, it runs $\mathcal{A}(10^{n-1}\|h(y))$ and obtains $x'\|y'$. It outputs y'.

Let us first analyze the running time of $\mathcal{B}$. The first two steps are clearly polynomial (in n) time. In the third step, $\mathcal{B}$ runs $\mathcal{A}$ and uses its output. Note that the running time of since $\mathcal{A}$ runs in polynomial (in n) time, this step also takes polynomial (in n) time. Thus, the overall running time of $\mathcal{B}$ is polynomial (in n).

Let us now calculate the probability that $\mathcal{B}$ outputs the correct inverse. If $x = 0^n$, the probability that y' is the correct inverse is at least $\frac{1}{2^n}$ (because it guesses y' randomly and probability that a random y' is the correct inverse is $\geq 1/2^n$). On the other hand, if $x \neq 0^n$, then the probability that $\mathcal{B}$ outputs the correct inverse is $\mu(n)$. Thus,

$$\begin{aligned} \Pr[\mathcal{B}(1^n, h(y)) \in h^{-1}(h(y))] &\geq \Pr[x = 0^n](\frac{1}{2^n}) + \Pr[x \neq 0^n]\mu(n) \\ &= \frac{1}{2^{2n}} + (1 - \frac{1}{2^n})\mu(n) \\ &\geq \mu(n) - (\frac{1}{2^n} - \frac{1}{2^{2n}}) \end{aligned}$$

Since $\mu(n)$ is a non-negligible function and $(\frac{1}{2^n} - \frac{1}{2^{2n}})$ is a negligible function, their difference is non-negligible.[3] This contradicts the one-wayness of h.

Step-2: Constructing g'. We construct the new function $g' : \{0,1\}^{2n} \to \{0,1\}^{2n-1}$ by dropping the first bit of g. That is,

$$g'(x\|y) = \begin{cases} 0^{n-1}\|y & \text{if } x = 0^n \\ 0^{n-1}\|g(y) & \text{otherwise} \end{cases}$$

[3]Exercise: Prove that if $\alpha(\cdot)$ is a non-negligible function and $\beta(\cdot)$ is a negligible function, then $(\alpha - \beta)(\cdot)$ is a non-negligible function.

Step-3: Inverting g'. We now want to prove that g' is not one-way. That is, we want to design an adversary $\mathcal{C}$ such that given 1^{2n} and $g'(x\|y)$ for a randomly chosen x, y, it outputs an element in the set $g^{-1}(g(x\|y)$. The description of $\mathcal{C}$ is as follows:

- On input 1^{2n} and $g'(x\|y)$, the adversary $\mathcal{C}$ parses $g'(x\|y)$ as $0^{n-1}\|\overline{y}$.
- It outputs $0^n\|\overline{y}$ as the inverse.

Notice that $g'(0^n\|\overline{y}) = 0^{n-1}\|\overline{y}$. Thus, $\mathcal{C}$ succeeds with probability 1 and this breaks the one-wayness of g'.

1.6 Hardness Amplification

In this section, we show that even a very *weak* form of one-way functions suffices from constructing one-way functions as defined previously. For this section, we refer to this previously defined notion as strong one-way functions.

Definition 1.6 (Weak One-Way Functions) *A function $f : \{0,1\}^n \to \{0,1\}^m$ is said to be a weak one-way function if:*

- *f is computable by a polynomial time machine, and*
- *There exists a noticeable function $\alpha_f(\cdot)$ such that $\forall$ non-uniform PPT adversaries $\mathcal{A}$ we have that*

$$\mu_{\mathcal{A},f}(n) = \Pr_{x \xleftarrow{\$} \{0,1\}^n}[\mathcal{A}(1^n, f(x)) \in f^{-1}(f(x))] \leq 1 - \alpha_f(n).$$

Theorem 1.1 *If there exists a weak one-way function, then there exists a (strong) one-way function.*

Proof. We prove the above theorem constructively. Suppose $f : \{0,1\}^n \to \{0,1\}^m$ is a weak one-way function, then we prove that the function $g : \{0,1\}^{nq} \to \{0,1\}^{mq}$ for $q = \lceil \frac{2n}{\alpha_f(n)} \rceil$

$$g(x_1, x_2, \cdots, x_q) = f(x_1)||f(x_2)|| \cdots ||f(x_q),$$

is a strong one-way function.

Assume for the sake of contradiction that there exists an adversary $\mathcal{B}$ such that $\mu_{\mathcal{B},g}(nq) = \Pr_{x \xleftarrow{\$} \{0,1\}^{nq}}[\mathcal{B}(1^{nq}, g(x)) \in g^{-1}(g(x))]$ is non-negligible. Then we use $\mathcal{B}$ to construct $\mathcal{A}$ (see Figure 1.1) that breaks f, namely $\mu_{\mathcal{A},f}(n) = \Pr_{x \xleftarrow{\$} \{0,1\}^n}[\mathcal{A}(1^n, f(x)) \in f^{-1}(f(x))] > 1 - \alpha_f(n)$ for sufficiently large n.

Note that: (1) $\mathcal{A}(1^n, y)$ iterates at most $T = \frac{4n^2}{\alpha_f(n)\mu_{\mathcal{B},g}(nq)}$ times each call is polynomial time. (2) $\mu_{\mathcal{B},g}(nq)$ is a non-negligible function. This implies that for infinite choices of n this value is greater than some noticeable function. Together these two facts imply that for infinite choices of n the running time of $\mathcal{A}$ is bounded by a polynomial function in n.

It remains to show that $\Pr_{x \xleftarrow{\$} \{0,1\}^n}[\mathcal{A}(1^n, f(x)) = \perp] < \alpha_f(n)$ for arbitrarily large n. A natural way to argue this is by showing that at least one execution of $\mathcal{B}$ should suffice for inverting $f(x)$.

loop $T = \frac{4n^2}{\alpha_f(n)\mu_{\mathcal{B},g}(nq)}$ times
  $i \xleftarrow{\$} \{1, 2, \cdots, q\}$
  $x_1, \cdots, x_{i-1}, x_i, \cdots, x_q \xleftarrow{\$} \{0,1\}^n$
  $(x'_1, x'_2, \cdots, x'_q) := \mathcal{B}(f(x_1), f(x_2), \cdots, f(x_q))$
  if $f(x'_i) = y$ **then**
    return x'_i
return $\perp$

Figure 1.1: Construction of $\mathcal{A}(1^n, y)$

However, the technical challenge in proving this formally is that these calls to $\mathcal{B}$ aren't independent. Below we formalize this argument even when these calls aren't independent.

Define the set S of "bad" x's, which are hard to invert:

$$S := \left\{ x \,\middle|\, \Pr_{\mathcal{B}}[\mathcal{A} \text{ inverts } f(x) \text{ in a single iteration}] \leq \frac{\alpha_f(n)\mu_{\mathcal{B},g}(nq)}{4n} \right\}.$$

We start by proving that the size of S is small. More formally,

$$\Pr_{x \xleftarrow{\$} \{0,1\}^n}[x \in S] \leq \frac{\alpha_f(n)}{2}.$$

Assume, for the sake of contradiction, that $\Pr_{x \xleftarrow{\$} \{0,1\}^n}[x \in S] > \frac{\alpha_f(n)}{2}$. Then we have that:

$$\begin{aligned}
\mu_{\mathcal{B},g}(nq) &= \Pr_{(x_1,\cdots,x_q) \xleftarrow{\$} \{0,1\}^{nq}}[\mathcal{B}(1^{nq}, g(x_1,\cdots,x_q)) \in g^{-1}(g(x_1,\cdots,x_q))] \\
&= \Pr_{x_1,\cdots,x_q}[\mathcal{B}(1^{nq}, g(x_1,\cdots,x_q)) \in g^{-1}(g(x_1,\cdots,x_q)) \wedge \forall i : x_i \notin S] \\
&\quad + \Pr_{x_1,\cdots,x_q}[\mathcal{B}(1^{nq}, g(x_1,\cdots,x_q)) \in g^{-1}(g(x_1,\cdots,x_q)) \wedge \exists i : x_i \in S] \\
&\leq \Pr_{x_1,\cdots,x_q}[\forall i : x_i \notin S] + \sum_{i=1}^{q} \Pr_{x_1,\cdots,x_q}[\mathcal{B}(1^{nq}, g(x_1,\cdots,x_q)) \in g^{-1}(g(x_1,\cdots,x_q)) \wedge x_i \in S] \\
&\leq \left(1 - \frac{\alpha_f(n)}{2}\right)^q + q \cdot \Pr_{x_1,\cdots,x_q,i}[\mathcal{B}(1^{nq}, g(x_1,\cdots,x_q)) \in g^{-1}(g(x_1,\cdots,x_q)) \wedge x_i \in S] \\
&= \left(1 - \frac{\alpha_f(n)}{2}\right)^{\frac{2n}{\alpha_f(n)}} + q \cdot \Pr_{x \xleftarrow{\$} \{0,1\}^n, \mathcal{B}}[\mathcal{A} \text{ inverts } f(x) \text{ in a single iteration} \wedge x \in S] \\
&\leq e^{-n} + q \cdot \Pr_x[x \in S] \cdot \Pr[\mathcal{A} \text{ inverts } f(x) \text{ in a single iteration} \mid x \in S] \\
&\leq e^{-n} + \frac{2n}{\alpha_f(n)} \cdot 1 \cdot \frac{\mu_{\mathcal{B},g}(nq) \cdot \alpha_f(n)}{4n} \\
&\leq e^{-n} + \frac{\mu_{\mathcal{B},g}(nq)}{2}.
\end{aligned}$$

Hence $\mu_{\mathcal{B},g}(nq) \leq 2e^{-n}$, contradicting with the fact that $\mu_{\mathcal{B},g}$ is non-negligible. Then we have

$$\begin{aligned}
&\Pr_{x \xleftarrow{\$} \{0,1\}^n} [\mathcal{A}(1^n, f(x)) = \bot] \\
&\quad = \Pr_x[x \in S] + \Pr_x[x \notin S] \cdot \Pr[\mathcal{B} \text{ fails to invert } f(x) \text{ in every iteration} | x \notin S] \\
&\quad \leq \frac{\alpha_f(n)}{2} + (\Pr[\mathcal{B} \text{ fails to invert } f(x) \text{ a single iteration} | x \notin S])^T \\
&\quad \leq \frac{\alpha_f(n)}{2} + \left(1 - \frac{\mu_{\mathcal{A},g}(nq) \cdot \alpha_f(n)}{4n}\right)^T \\
&\quad \leq \frac{\alpha_f(n)}{2} + e^{-n} \leq \alpha_f(n)
\end{aligned}$$

for sufficiently large n. This concludes the proof. ■

1.7 Levin's One-Way Function

Theorem 1.2 *If there exists a one-way function, then there exists an explicit function f that is one-way (constructively).*

Lemma 1.2 *If there exists a one-way function computable in time n^c for a constant c, then there exists a one-way function computable in time n^2.*

Proof. Let $f : \{0,1\}^n \to \{0,1\}^n$ be a one-way function computable in time n^c. Construct $g : \{0,1\}^{n+n^c} \to \{0,1\}^{n+n^c}$ as follows:

$$g(x, y) = f(x) || y$$

where $x \in \{0,1\}^n, y \in \{0,1\}^{n^c}$. $g(x, y)$ takes time $2n^c$, which is linear in the input length.

We next show that $g(\cdot)$ is one-way. Assume for the purpose of contradiction that there exists an adversary $\mathcal{A}$ such that $\mu_{\mathcal{A},g}(n + n^c) = \Pr_{(x,y) \xleftarrow{\$} \{0,1\}^{n+n^c}} [\mathcal{A}(1^{n+n^c}, g(x, y)) \in g^{-1}(g(x, y))]$ is non-negligible. Then we use $\mathcal{A}$ to construct $\mathcal{B}$ such that $\mu_{\mathcal{B},f}(n) = \Pr_{x \xleftarrow{\$} \{0,1\}^n} [\mathcal{B}(1^n, f(x)) \in f^{-1}(f(x))]$ is also non-negligible.

$\mathcal{B}$ on input $z \in \{0,1\}^n$, samples $y \xleftarrow{\$} \{0,1\}^{n^c}$, and outputs the n higher-order bits of $\mathcal{A}(1^{n+n^c}, z||y)$. Then we have

$$\begin{aligned}
\mu_{\mathcal{B},g}(n) = &\Pr_{x \xleftarrow{\$} \{0,1\}^n, y \xleftarrow{\$} \{0,1\}^{n^c}} \left[\mathcal{A}(1^{n+n^c}, f(x)||y) \in f^{-1}(f(x))||\{0,1\}^{n^c}\right] \\
\geq &\Pr_{x,y} \left[\mathcal{A}(1^{n+n^c}, g(x, y)) \in f^{-1}(f(x))||y\right] \\
= &\Pr_{x,y} \left[\mathcal{A}(1^{n+n^c}, g(x, y)) \in g^{-1}(g(x, y))\right]
\end{aligned}$$

is non-negligible. ■

Proof. [of Theorem 1.2] We first construct a weak one-way function $h : \{0,1\}^n \to \{0,1\}^n$ as follows:

$$h(M, x) = \begin{cases} M||M(x) & \text{if } M(x) \text{ takes no more than } |x|^2 \text{ steps} \\ M||0 & \text{otherwise} \end{cases}$$

where $|M| = \log n, |x| = n - \log n$ (interpreting M as the code of a machine and x as its input). If h is weak one-way, then we can construct a one-way function from h as we discussed in Section 1.6.

It remains to show that if one-way functions exist, then h is a weak one-way function, with $\alpha_h(n) = \frac{1}{n^2}$. Assume for the purpose of contradiction that there exists an adversary $\mathcal{A}$ such that $\mu_{\mathcal{A},h}(n) = \Pr_{(M,x)\xleftarrow{\$}\{0,1\}^n}[\mathcal{A}(1^n, h(M,x)) \in h^{-1}(h(M,x))] \geq 1 - \frac{1}{n^2}$ for all sufficiently large n. By the existence of one-way functions and Lemma 1.2, there exists a one-way function $\tilde{M}$ that can be computed in time n^2. Let $\tilde{M}$ be the uniform machine that computes this one-way function. We will consider values n such that $n > 2^{|\tilde{M}|}$. In other words for these choices of n, $\tilde{M}$ can be described using $\log n$ bits. We construct $\mathcal{B}$ to invert $\tilde{M}$: on input y outputs the $(n - \log n)$ lower-order bits of $\mathcal{A}(1^n, \tilde{M}||y)$. Then

$$\begin{aligned}\mu_{\mathcal{B},\tilde{M}}(n - \log n) &= \Pr_{x\xleftarrow{\$}\{0,1\}^{n-\log n}}\left[\mathcal{A}(1^n, \tilde{M}||\tilde{M}(x)) \in \{0,1\}^{\log n}||\tilde{M}^{-1}(\tilde{M}((x))\right] \\ &\geq \Pr_{x\xleftarrow{\$}\{0,1\}^{n-\log n}}\left[\mathcal{A}(1^n, \tilde{M}||\tilde{M}(x)) \in \tilde{M}||\tilde{M}^{-1}(\tilde{M}((x))\right].\end{aligned}$$

Observe that for sufficiently large n it holds that

$$\begin{aligned}1 - \frac{1}{n^2} \leq& \mu_{\mathcal{A},h}(n) \\ =& \Pr_{(M,x)\xleftarrow{\$}\{0,1\}^n}\left[\mathcal{A}(1^n, h(M,x)) \in h^{-1}(h(M,x))\right] \\ \leq& \Pr_M[M = \tilde{M}] \cdot \Pr_x\left[\mathcal{A}(1^n, \tilde{M}||\tilde{M}(x)) \in \tilde{M}||\tilde{M}^{-1}(\tilde{M}((x))\right] + \Pr_M[M \neq \tilde{M}] \\ \leq& \frac{1}{n} \cdot \mu_{\mathcal{B},\tilde{M}}(n - \log n) + \frac{n-1}{n}.\end{aligned}$$

Hence $\mu_{\mathcal{B},\tilde{M}}(n - \log n) \geq \frac{n-1}{n}$ for sufficiently large n which is a contradiction. ■

Exercises

Exercise 1.1 *If $\mu(\cdot)$ and $\nu(\cdot)$ are negligible functions then show that $\mu(\cdot) \cdot \nu(\cdot)$ is a negligible function.*

Exercise 1.2 *If $\mu(\cdot)$ is a negligible function and $f(\cdot)$ is a function polynomial in its input then show that $\mu(f(\cdot))$[4] are negligible functions.*

Exercise 1.3 *Prove that the existence of one-way functions implies $P \neq NP$.*

Exercise 1.4 *Prove that there is no one-way function $f : \{0,1\}^n \to \{0,1\}^{\lfloor \log_2 n \rfloor}$.*

Exercise 1.5 *Let $f : \{0,1\}^n \to \{0,1\}^n$ be any one-way function then is $f'(x) \stackrel{def}{=} f(x) \oplus x$ necessarily one-way?*

Exercise 1.6 *Prove or disprove: If $f : \{0,1\}^n \to \{0,1\}^n$ is a one-way function, then $g : \{0,1\}^n \to \{0,1\}^{n-\log n}$ is a one-way function, where $g(x)$ outputs the $n - \log n$ higher order bits of $f(x)$.*

Exercise 1.7 *Explain why the proof of Theorem 1.1 fails if the attacker $\mathcal{A}$ in Figure 1.1 sets $i = 1$ and not $i \xleftarrow{\$} \{1, 2, \cdots, q\}$.*

Exercise 1.8 *Given a (strong) one-way function construct a weak one-way function that is not a (strong) one-way function.*

Exercise 1.9 *Let $f : \{0,1\}^n \to \{0,1\}^n$ be a weak one-way permutation (a weak one way function that is a bijection). More formally, f is a PPT computable one-to-one function such that $\exists$ a constant $c > 0$ such that $\forall$ non-uniform PPT machine A and $\forall$ sufficiently large n we have that:*

$$\Pr_{x,A}[A(f(x)) \notin f^{-1}(f(x))] > \frac{1}{n^c}$$

Show that $g(x) = f^T(x)$ is not a strong one way permutation. Here f^T denotes the T times self composition of f and T is a polynomial in n.

Interesting follow up reading if interested: With some tweaks the function above can be made a strong one-way permutation using explicit constructions of expander graphs. See Section 2.6 in `http://www.wisdom.weizmann.ac.il/~oded/PSBookFrag/part2N.ps`

[4] Assume that μ and f are such that $\mu(f(\cdot))$ takes inputs from $\mathbb{Z}^+$ and outputs values in $[0, 1]$.