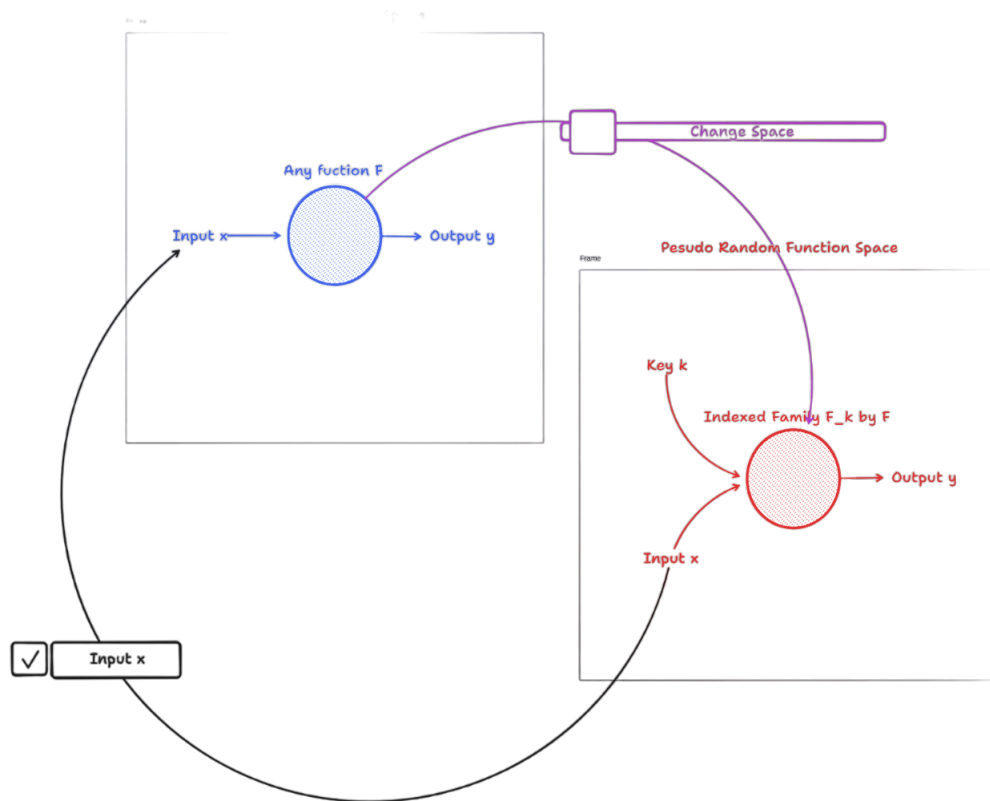


VisualCrypt: Essence of Provable Security

Ji Yong-Hyeon, Kim Dong-Hyeon



Department of Information Security, Cryptology, and Mathematics
College of Science and Technology
Kookmin University

December 14, 2023

VISUALCRYPT: ESSENCE OF PROVABLE SECURITY
DEPARTMENT OF INFORMATION SECURITY, CRYPTOLOGY, AND MATHEMATICS
THE COLLEGE OF SCIENCE AND TECHNOLOGY
KOOKMIN UNIVERSITY

No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the author, except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law. For permission requests, write to the author at the address below.

hacker3740@gmail.com

Copyright © 2023, Ji Yong-Hyeon. All rights reserved.

Contents

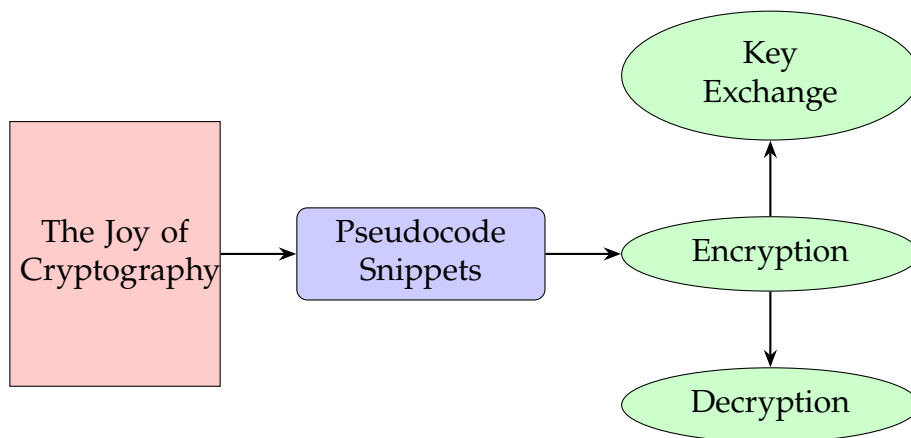
Introduction	2
1 One-Time Pad & Kerckhoff's Principle	3
2 The Basics of Provable Security	5
2.1 How to Write a Security Definition	5
2.1.1 Syntax and Correctness	5
2.1.2 "Real-vs-Random" Style of Security Definition	6
2.2 Formalisms for Security Definition	7
3 Cryptography on Intractable Computations	9
3.1 What Qualifies as a "Computationally Infeasible" Attack?	9
3.2 What Qualifies as a "Negligible" Success Probability?	10
3.3 Indistinguishability	12
4 Pseudo-random Generators (PRG)	20
4.1 Definition	20
4.2 Shorter Keys in One-Time-Secret Encryption	24
5 Pseudo-Random Functions & Block Ciphers	26
5.1 Definition	27
6 Security Against Chosen Plaintext Attacks	31
6.1 Limits of Deterministic Encryption	32

List of Symbols

λ	Security Parameter
$\mathbf{0}^\lambda, \mathbf{1}^\lambda$	$\underbrace{\mathbf{00} \cdots \mathbf{0}}_{\lambda \text{ times}}, \underbrace{\mathbf{11} \cdots \mathbf{1}}_{\lambda \text{ times}} : \lambda\text{-bit zero/one sequence}$
$\{\mathbf{0}, \mathbf{1}\}$	Binary Field
\mathcal{L}	Library
\mathcal{A}	Adversary
$\mathcal{A} \diamond \mathcal{L}$	The result of linking \mathcal{A} to \mathcal{L}
$\$$	Randomness
$\xleftarrow{\$}$	Uniformly Chosen
\equiv	Interchangability; Identical
\approx	Indistinguishability Symbol
\mathcal{K}	Key Space
\mathcal{M}	Message Space
\mathcal{C}	Ciphertext Space
$\text{CTXT}()$	Ciphertext Output Function
$\text{EAVESDROP}()$	Eavesdrop Function

Introduction

“The Joy of Cryptography” is a unique book that looks at cryptographic "security proofs" from a coding perspective. Instead of using complex mathematics, it uses pseudocode. Pseudocode is like a simple way to write down instructions for a computer. This approach is thought to bring a fresh viewpoint on how we prove the security of cryptographic methods.



Inspired by this book, we have also tried to develop cryptographic security proofs using pseudocode. We have taken these pseudocode examples and used various coding techniques to create visualizations. These visualizations turn the pseudocode into pictures and diagrams. This makes it easier to see and understand how cryptographic security works.

We believe that this method will make it easier for more people to understand and get involved in cryptographic security certification. By using pseudocode and visualizations, we hope to lower the barrier that often makes cryptography seem difficult and inaccessible.

Chapter 1

One-Time Pad & Kerckhoff's Principle

Kerckhoffs' Principle:

Design your system to be secure even if the attacker has complete knowledge of all its algorithms.

One-time Pad (OTP)

Construction 1.1. The specific KeyGen, Enc, and Dec algorithms for **one-time pad** are given below:

KeyGen :	Enc($k, m \in \{0, 1\}^\lambda$) :	Dec($k, c \in \{0, 1\}^\lambda$) :
$k \xleftarrow{\$} \{0, 1\}^\lambda$ return k	return $k \oplus m$	return $k \oplus c$

Algorithm 1: Key Generation, Encryption, and Decryption

```
1: Function KeyGen():
2:    $k \leftarrow \{0, 1\}^\lambda$ 
3:   return  $k$ 
4: return

5: Function Enc( $k, m \in \{0, 1\}^\lambda$ ):
6:   return  $k \oplus m$ 
7: return

8: Function Dec( $k, c \in \{0, 1\}^\lambda$ ):
9:   return  $k \oplus c$ 
10: return
```

Correctness of OTP

Proposition 1.1.

$$(\forall k, m \in \{\mathbf{0}, \mathbf{1}\}^\lambda) \quad \text{Dec}(k, \text{Enc}(k, m)) = m.$$

Proof. Let $k, m \in \{\mathbf{0}, \mathbf{1}\}^\lambda$ then

$$\begin{aligned} \text{Dec}(k, \text{Enc}(k, m)) &= \text{Dec}(k, k \oplus m) = k \oplus (k \oplus m) \\ &= (k \oplus k) \oplus m \\ &= \mathbf{0}^\lambda \oplus m \\ &= m. \end{aligned}$$

□

Remark 1.1 (EAVESDROP Algorithm). From Eve's perspective, seeing a ciphertext corresponds to receiving an output from the following algorithm:

$\begin{array}{l} \text{EAVESDROP}(m \in \{\mathbf{0}, \mathbf{1}\}^\lambda) \\ \hline k \xleftarrow{\$} \{\mathbf{0}, \mathbf{1}\}^\lambda \\ c := k \oplus m \\ \text{return } c \end{array}$

Theorem 1.2. Let $m \in \{\mathbf{0}, \mathbf{1}\}^\lambda$. The distribution $\text{EAVESDROP}(m)$ is the **uniform distribution** on $\{\mathbf{0}, \mathbf{1}\}^\lambda$. In other words,

$$m, m' \in \{\mathbf{0}, \mathbf{1}\}^\lambda \implies \text{dist}(\text{EAVESDROP}(m)) \sim \text{dist}(\text{EAVESDROP}(m')).$$

Chapter 2

The Basics of Provable Security

2.1 How to Write a Security Definition

2.1.1 Syntax and Correctness

Encryption Syntax

Definition 2.1. A **symmetric-key encryption (SKE) scheme** consists of the following algorithms:

- KeyGen outputs a key $k = \text{KeyGen}(1^\lambda) \in \mathcal{K}$
- $\text{Enc} : \mathcal{K} \times \mathcal{M} \rightarrow \mathcal{C}$
- $\text{Dec} : \mathcal{K} \times \mathcal{C} \rightarrow \mathcal{M}$

We call \mathcal{K} the **key space**, \mathcal{M} the **message space**, and \mathcal{C} the **ciphertext space** of the scheme.

Remark 2.1. Note that

- KeyGen is a randomized algorithm¹.
- Enc is a (possibly randomized) algorithm².
- Dec is a deterministic algorithm³.

Remark 2.2. We refer to the entire scheme by a single variable Σ , i.e.,

$$\Sigma = (\text{KeyGen}, \text{Enc}, \text{Dec}).$$

Remark 2.3. We write

$$\begin{array}{ccc} \Sigma.\text{KeyGen}, & \Sigma.\text{Enc}, & \Sigma.\text{Dec}, \\ \Sigma.\mathcal{K}, & \Sigma.\mathcal{M}, & \Sigma.\mathcal{C} \end{array}$$

to refer to its components.

¹An algorithm that makes use of random numbers.

²It could operate deterministically or non-deterministically depending on specific conditions or parameters.

³An algorithm that does produces the same output for the same input, every time it's run.

SKE Correctness

Definition 2.2. An encryption scheme Σ satisfies **correctness** if

$$(\forall k \in \Sigma.\mathcal{K}) (\forall m \in \Sigma.\mathcal{M}) \quad \Pr [\Sigma.\text{Dec}(k, \Sigma.\text{Enc}(k, m)) = m] = 1.$$

Remark 2.4. The definition is written in terms of a probability because Enc is allowed to be a randomized algorithm. In other words, decrypting a ciphertext with the same key that was used for encryption must *always* result in the original plaintext.

2.1.2 “Real-vs-Random” Style of Security Definition

“an encryption scheme is a good one if its ciphertexts *look like* random junk to an attacker”

Security definitions always consider **the attacker’s view** of the system.

“an encryption scheme is a good one if its ciphertexts *look like* random junk to an attacker ... when each key is secret and used to encrypt only one plaintext, even when the attacker chooses the plaintexts.”

A concise way to express all of these details is to consider **the attacker as a calling program** to the following subroutine:

$\begin{array}{l} \text{CTXT}(m \in \Sigma.\mathcal{M}) : \\ \hline k \leftarrow \Sigma.\text{KeyGen} \\ c := \Sigma.\text{Enc}(k, m) \\ \text{return } c \end{array}$
--

Algorithm 2: CTXT function

```

1: Function CTXT( $m \in \Sigma.\mathcal{M}$ ):
2:    $k \leftarrow \Sigma.\text{KeyGen}$ 
3:    $c := \Sigma.\text{Enc}(k, m)$ 
4:   return  $c$ 
5: end

```

Example 2.1 (One-Time Pad (OTP)). a

$\begin{array}{l} \text{CTXT}(m) : \\ \hline k \leftarrow \{\mathbf{0}, \mathbf{1}\}^\lambda \quad // \text{KeyGen of OTP} \\ c := k \oplus m \quad // \text{Enc of OTP} \\ \text{return } c \end{array}$

vs.

$\begin{array}{l} \text{CTXT}(m) : \\ \hline c := \{\mathbf{0}, \mathbf{1}\}^\lambda \quad // C \text{ of OTP} \\ \text{return } c \end{array}$

“an encryption scheme is a good one if, when you plug its KeyGen and Enc algorithms into the template of the CTXT subroutine above, the two implementations of CTXT induce identical behavior in every calling program.”

2.2 Formalisms for Security Definition

Library

Definition 2.3. A library \mathcal{L} is a collection of subroutines and private/static variables.

Example 2.2. Here is a familiar library and one possible calling program:

\mathcal{L}		\mathcal{A}
$\text{CTXT}(m):$ $k \xleftarrow{\$} \{0, 1\}^\lambda$ $c := k \oplus m$ $\text{return } c$,	$m \leftarrow \{0, 1\}^\lambda$ $c := \text{CTXT}(m)$ $\text{return } m \stackrel{?}{=} c$

Then

$$\Pr[\mathcal{A} \diamond \mathcal{L} \Rightarrow \text{true}] = \frac{1}{2^\lambda}.$$

Algorithm 3: Adversary \mathcal{A}

```

1: Function Adversary  $\mathcal{A}$ :
2:    $m \leftarrow \{0, 1\}^\lambda$ 
3:    $c := \text{CTXT}(m)$ 
4:   return  $m = c$ 
5: end

```

Interchangeability

Definition 2.4. Let \mathcal{L}_1 and \mathcal{L}_2 be two libraries that have the same interface. We say that \mathcal{L}_1 and \mathcal{L}_2 are **interchangeable**, and write $\mathcal{L}_1 \equiv \mathcal{L}_2$, if $\forall \mathcal{A}$:

$$\Pr[\mathcal{A} \diamond \mathcal{L}_1 \Rightarrow \text{true}] = \Pr[\mathcal{A} \diamond \mathcal{L}_2 \Rightarrow \text{true}].$$

One-Time Uniform Ctxts

Definition 2.5. An encryption scheme Σ has **one-time uniform cipher-texts** if

$\mathcal{L}_{\text{ots\$-real}}^\Sigma$		$\mathcal{L}_{\text{ots\$-rand}}^\Sigma$
$\text{CTXT}(m \in \Sigma.\mathcal{M}):$ $k \leftarrow \Sigma.\text{KeyGen}$ $c \leftarrow \Sigma.\text{Enc}(k, m)$ $\text{return } c$	\equiv	$\text{CTXT}(m \in \Sigma.\mathcal{M}):$ $c \leftarrow \Sigma.C$ $\text{return } c$

One-Time Secrecy (OTS)

Definition 2.6. One-time secrecy is a property of an encryption scheme where an adversary cannot gain any information about the plaintext message from the ciphertext, even if they know the encryption key was used only once.

$\mathcal{L}_{\text{ots-L}}^\Sigma$ <hr style="border: 0; border-top: 1px solid black; margin: 5px 0;"/> $\text{Eve}(m_L, m_R \in \Sigma.\mathcal{M}):$ $k \leftarrow \Sigma.\text{KeyGen}$ $c \leftarrow \Sigma.\text{Enc}(k, m_L)$ return c	\equiv	$\mathcal{L}_{\text{ots-R}}^\Sigma$ <hr style="border: 0; border-top: 1px solid black; margin: 5px 0;"/> $\text{Eve}(m_L, m_R \in \Sigma.\mathcal{M}):$ $k \leftarrow \Sigma.\text{KeyGen}$ $c \leftarrow \Sigma.\text{Enc}(k, m_R)$ return c
---	----------	---

Chapter 3

Cryptography on Intractable Computations

3.1 What Qualifies as a “Computationally Infeasible” Attack?

Polynomial Time

Definition 3.1. A program runs in **polynomial time** if

$$\exists c > 0 : \forall n \geq n_0 : \text{Time}(n) \leq n^c,$$

where Time is the time taken by the algorithm on inputs of size n . n_0 is constant size of the input. That is, there exists a constant $c > 0$ such that for all sufficiently long input strings x with $|x| = n$, the program stops after no more than $O(n^c)$ steps.

Remark 3.1. We see “polynomial-time” as a synonym for “efficient.”

Example 3.1. $\text{gcd}(a, b)$ can be computed using $O((\log_2 a)^3)$ bit operation if $a > b$.

Example 3.2.

Efficient algorithm known:	No known efficient algorithm:
Computing GCDs	Factoring integers
Arithmetic mod N	Computing $\phi(N)$ given N
Inverses mod N	Discrete logarithm
Exponentiation mod N	Square roots mod composite N

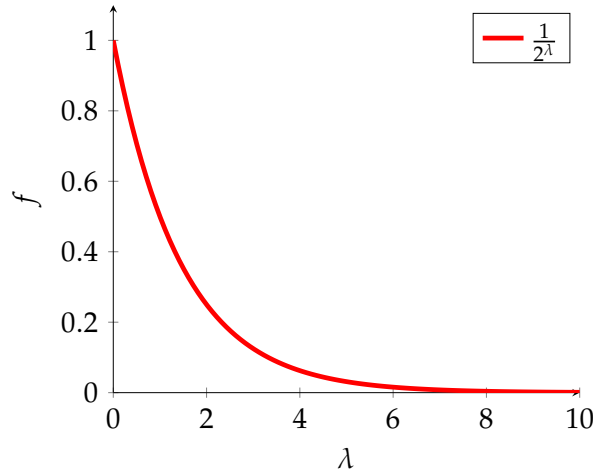
Again, “efficient” means polynomial-time. Furthermore, we only consider polynomial-time algorithms that run on standard, *classical* computers. In fact, all of the problems in the right-hand column *do* have known polynomial-time algorithms on *quantum* computers.

3.2 What Qualifies as a “Negligible” Success Probability?

For a cryptographic system to be considered secure, we often want the success probability of any polynomial-time adversary to be negligible in the security parameter λ .

Idea. $\frac{1}{2^\lambda}$ approaches zero so fast that no polynomial can “rescue”.

Proof. Assume that $f(\lambda) = \frac{1}{2^\lambda}$.



Consider any polynomial $p(\lambda)$ of degree n , written as:

$$p(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots + a_n\lambda^n = \sum_{i=0}^n a_i\lambda^i.$$

The product $p(\lambda)$ and $f(\lambda)$ is

$$p(\lambda)f(\lambda) = a_0\frac{1}{2^\lambda} + a_1\frac{\lambda}{2^\lambda} + \cdots + a_n\frac{\lambda^n}{2^\lambda}.$$

We claim that $\lim_{\lambda \rightarrow \infty} a_k \frac{\lambda^k}{2^\lambda} = 0$, where $k \in \mathbb{Z}_{\geq 0}$. Let $g(\lambda) = \lambda^k$ and $h(\lambda) = 2^\lambda$. Note that

$$\begin{aligned} h'(\lambda) &= 2^\lambda(\ln 2), & g'(\lambda) &= k\lambda^{k-1} \\ h''(\lambda) &= 2^\lambda(\ln 2)^2, & g''(\lambda) &= k(k-1)\lambda^{k-2} \\ &\vdots & & \\ h^{(k)}(\lambda) &= 2^\lambda(\ln 2)^k, & g^{(k)}(\lambda) &= k!. \end{aligned}$$

By applying L'Hôpital's Rule k times, we have

$$\lim_{\lambda \rightarrow \infty} \frac{\lambda^k}{2^\lambda} = \lim_{\lambda \rightarrow \infty} \frac{k!}{2^\lambda(\ln 2)^k} = 0.$$

Thus, $\lim_{\lambda \rightarrow \infty} p(\lambda)f(\lambda) = 0$. □

Negligible

Definition 3.2. A function f is **negligible** if,

$$\forall \text{polynomial } p : \lim_{\lambda \rightarrow \infty} p(\lambda)f(\lambda) = 0.$$

In other words, a negligible function approaches zero so fast that you can never catch up when multiplying by a polynomial.

Remark 3.2. As λ (security parameter) gets larger and larger, the product of $p(\lambda)$ (resources or capabilities for an adversary) and $f(\lambda)$ (success probability) approaches 0.

Remark 3.3. A function $f(\lambda)$ is negligible if $\forall p(\lambda) > 0 : \exists \lambda_0 : \lambda > \lambda_0 \Rightarrow |f(\lambda)| < \frac{1}{p(\lambda)}$.

Proposition 3.1. Let $c \in \mathbb{Z}$.

$$\lim_{\lambda \rightarrow \infty} \lambda^c f(\lambda) = 0 \implies f \text{ is negligible.}$$

Proof. Suppose that f satisfies $\lim_{\lambda \rightarrow \infty} \lambda^c f(\lambda) = 0$ for any $c \in \mathbb{Z}$, and take an arbitrary polynomial p of degree n . Since $\lim_{\lambda \rightarrow \infty} \frac{p(\lambda)}{\lambda^{n+1}} = 0$, we have

$$\lim_{\lambda \rightarrow \infty} p(\lambda)f(\lambda) = \lim_{\lambda \rightarrow \infty} \left[\frac{p(\lambda)}{\lambda^{n+1}} \left(\lambda^{n+1} \cdot f(\lambda) \right) \right] = \left(\lim_{\lambda \rightarrow \infty} \frac{p(\lambda)}{\lambda^{n+1}} \right) \left(\lim_{\lambda \rightarrow \infty} \lambda^{n+1} f(\lambda) \right) = 0 \cdot 0 = 0.$$

□

Example 3.3. Let $c \in \mathbb{Z}$. Then

$$\lim_{\lambda \rightarrow \infty} \lambda^c \frac{1}{2^\lambda} = \lim_{\lambda \rightarrow \infty} \frac{(\lambda^c)^{\log_2 2}}{2^\lambda} = \lim_{\lambda \rightarrow \infty} \frac{2^{c \log_2 \lambda}}{2^\lambda} = \lim_{\lambda \rightarrow \infty} 2^{c \log_2(\lambda) - \lambda} = 0$$

since $c \log_2(\lambda) - \lambda \rightarrow -\infty$ as $\lambda \rightarrow \infty$. Thus, $1/2^\lambda$ is negligible.

 $f \approx g$

Definition 3.3. Let $f, g : \mathbb{N} \rightarrow \mathbb{R}$ are real-valued functions. We write $f \approx g$ to mean that $|f(\lambda) - g(\lambda)|$ is a negligible function.

Remark 3.4. We use the terminology of negligible functions exclusively when discussing probabilities, so the following are common:

$\Pr[X] \approx 0 \Leftrightarrow$ “event X almost never happens”

$\Pr[Y] \approx 1 \Leftrightarrow$ “event Y almost always happens”

$\Pr[A] \approx \Pr[B] \Leftrightarrow$ “event A and B happen with essentially the same probability”

Additionally, the \approx symbol is *transitive*:

$$\Pr[X] \approx \Pr[Y] \wedge \Pr[Y] \approx \Pr[Z] \implies \Pr[X] \approx \Pr[Z].$$

3.3 Indistinguishability

Indistinguishable (\approx)

Definition 3.4. Let \mathcal{L}_1 and \mathcal{L}_2 be two libraries with a common interface, and let \mathcal{A} is a polynomial-time program that output a single bit. We say that \mathcal{L}_1 and \mathcal{L}_2 are **indistinguishable**, and write $\mathcal{L}_1 \approx \mathcal{L}_2$, if

$$\Pr[\mathcal{A} \diamond \mathcal{L}_1 \Rightarrow 1] \approx \Pr[\mathcal{A} \diamond \mathcal{L}_2 \Rightarrow 1].$$

Remark 3.5.

(1) We call the quantity

$$|\Pr[\mathcal{A} \diamond \mathcal{L}_1 \Rightarrow 1] - \Pr[\mathcal{A} \diamond \mathcal{L}_2 \Rightarrow 1]|$$

the **advantage** (or **bias**) of \mathcal{A} in distinguishing \mathcal{L}_1 and \mathcal{L}_2 .

(2) Two libraries are indistinguishable if all polynomial-time calling programs have negligible advantage in distinguishing them.

Example 3.4. Two indistinguishable libraries:

\mathcal{L}_1	\mathcal{L}_2
Predict(x): $s \xleftarrow{\$} \{0, 1\}^\lambda$ return $x \stackrel{?}{=} s$	Predict(x): return false

Algorithm 4: Algorithm \mathcal{L}_1 and \mathcal{L}_2

```

1  $\mathcal{L}_1$  Predict(x):
2    $s \leftarrow \{0, 1\}^\lambda$ 
3   return  $x = s$ 
4 end
5  $\mathcal{L}_2$  Predict(x):
6   return false
7 end

```

The calling program \mathcal{A} repeatedly invokes the ‘Predict’ functions and returns ‘1’ if it ever obtains a ‘true’ value from the response:

\mathcal{A}
do q times: if $\text{Predict}(0^\lambda) = \text{true}$ return 1 return 0

Algorithm 5: Adversarial Algorithm \mathcal{A}

```

1   $\mathcal{A}$ 
2  |   for  $i \leftarrow 1$  to  $q$  do
3  |   |   if  $\text{Predict}(\mathcal{O}^\lambda) = \text{true}$  then
4  |   |   |   return 1
5  |   |   end
6  |   end
7  |   return 0
8  end

```

(1) \mathcal{L}_2 can never return true, i.e., $\Pr[\mathcal{A} \diamond \mathcal{L}_2 \Rightarrow 1] = 0$.

(2) $\Pr[\mathcal{A} \diamond \mathcal{L}_2 \Rightarrow 1]$ is surely non-zero.

$$\Pr[\mathcal{A} \diamond \Rightarrow 1] = 1 - \Pr[\mathcal{A} \diamond \mathcal{L}_1 \Rightarrow 0] = 1 - \left(1 - \frac{1}{2^\lambda}\right)^q.$$

Using the union bound, we get:

$$\begin{aligned}
\Pr[\mathcal{A} \diamond \Rightarrow 1] &\leq \Pr[\text{first call to Predict returns true}] \\
&\quad + \Pr[\text{second call to Predict returns true}] \\
&\quad + \dots \\
&= q \cdot \frac{1}{2^\lambda}.
\end{aligned}$$

We showed that \mathcal{A} has non-zero advantage, and so $\mathcal{L}_1 \neq \mathcal{L}_2$. We also showed that \mathcal{A} has advantage at most $q/2^\lambda$. Since \mathcal{A} runs in polynomial time, it can only make a polynomial number q of queries to the library, so $q/2^\lambda$ is negligible.

Lemma 3.2.

(1) $\mathcal{L}_1 \equiv \mathcal{L}_2 \implies \mathcal{L}_1 \approx \mathcal{L}_2$.

(2) $\mathcal{L}_1 \approx \mathcal{L}_2 \approx \mathcal{L}_3 \implies \mathcal{L}_1 \approx \mathcal{L}_3$.

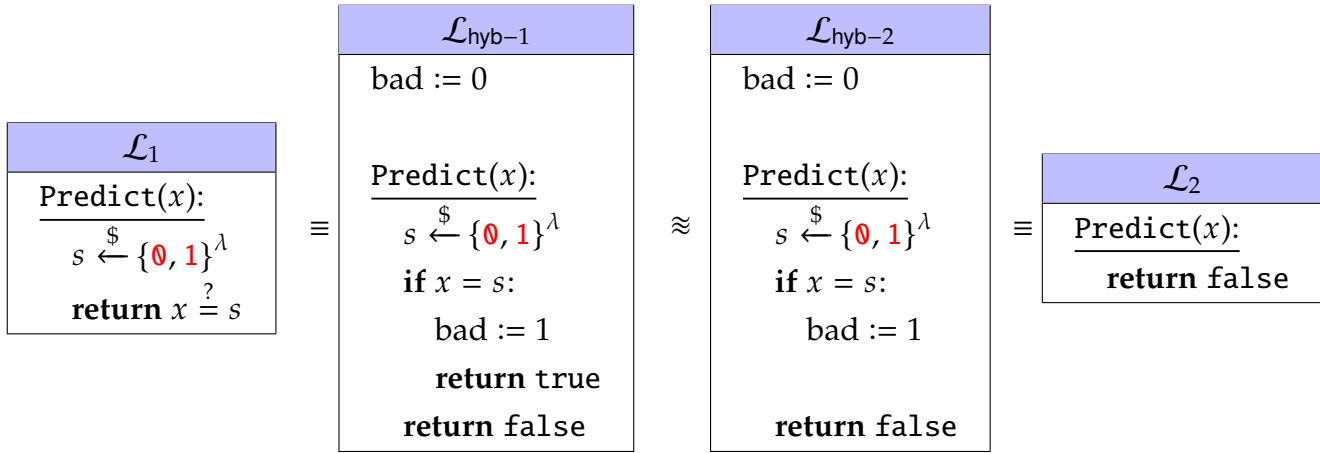
Lemma 3.3. For any polynomial-time library \mathcal{L}^* ,

$$\mathcal{L}_1 \approx \mathcal{L}_2 \implies \mathcal{L}^* \diamond \mathcal{L}_1 \approx \mathcal{L}^* \diamond \mathcal{L}_2.$$

Bad-Event Lemma**Lemma 3.4.**

$$|\Pr[\mathcal{A} \diamond \mathcal{L}_1 \Rightarrow 1] - \Pr[\mathcal{A} \diamond \mathcal{L}_2 \Rightarrow 1]| \leq \Pr[\mathcal{A} \diamond \mathcal{L}_1 \text{ sets bad} = 1].$$

Example 3.5. Consider \mathcal{L}_1 and \mathcal{L}_2 . They are indistinguishable with the following sequence of hybrids:



Algorithm 6: Algorithm \mathcal{L}_1 and $\mathcal{L}_{\text{hyb-1}}$

```

1   $\mathcal{L}_1$  Predict(x):
2  |    $s \leftarrow \{\mathbf{0}, \mathbf{1}\}^\lambda$ 
3  |   return  $x = s$ 
4  end
5   $\mathcal{L}_{\text{hyb-1}}$ 
6  |    $\text{bad} := 0$ 
7  |   Function Predict(x):
8  |   |    $s \leftarrow \{\mathbf{0}, \mathbf{1}\}^\lambda$ 
9  |   |   if  $x = s$  then  $\text{bad} := 1$ 
10 |   |   return true
11 |   |   return false
12 |   return
13 end

```

Algorithm 7: Algorithm $\mathcal{L}_{\text{hyb-2}}$ and \mathcal{L}_2

```

1   $\mathcal{L}_2$  Predict(x):
2  |   return false
3  end
4   $\mathcal{L}_{\text{hyb-2}}$ 
5  |    $\text{bad} := 0$ 
6  |   Function Predict(x):
7  |   |    $s \leftarrow \{\mathbf{0}, \mathbf{1}\}^\lambda$ 
8  |   |   if  $x = s$  then  $\text{bad} := 1$ 
9  |   |   return false
10 |   return
11 end

```

- $\mathcal{L}_1 \equiv \mathcal{L}_{\text{hyb-1}}$; Without *accessing* the variable “bad”, the change can have no effect.
- $\mathcal{L}_{\text{hyb-1}} \approx \mathcal{L}_{\text{hyb-2}}$; By the bad-event lemma,

$$\left| \Pr [\mathcal{A} \diamond \mathcal{L}_{\text{hyb-1}} \Rightarrow 1] - \Pr [\mathcal{A} \diamond \mathcal{L}_{\text{hyb-2}} \Rightarrow 1] \right| \leq \Pr [\mathcal{A} \diamond \mathcal{L}_{\text{hyb-1}} \text{ sets bad} = 1] .$$

- $\mathcal{L}_{\text{hyb-2}} \equiv \mathcal{L}_2$; Regardless of input, the subroutine always returns false.

Hence

$$\mathcal{L}_1 \equiv \mathcal{L}_{\text{hyb-1}} \approx \mathcal{L}_{\text{hyb-2}} \equiv \mathcal{L}_2 \implies \mathcal{L}_1 \approx \mathcal{L}_2 .$$

Exercises

4.2. The following are negligible functions in λ ? Justify your answers.

$$\frac{1}{\lambda^{1/\lambda}}.$$

Solution. $\frac{1}{\lambda^{1/\lambda}}$ is non-negligible functions.

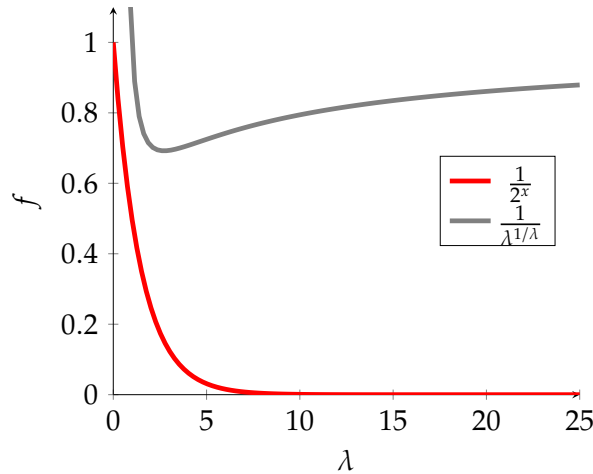


Figure 3.1: Non-negligible functions.

□

4.4. Show that when f is negligible, then for every polynomial p , the function $p(\lambda)f(\lambda)$ not only approaches 0, but it is also negligible itself.

Solution. We want to show that

$$p(\lambda)f(\lambda) \text{ is non-negligible} \implies f \text{ is non-negligible.}$$

Suppose that

$$\exists \text{ polynomial } q(\lambda) : \lim_{\lambda \rightarrow \infty} q(\lambda)p(\lambda)f(\lambda) = c \neq 0.$$

Then p is non-zero polynomial and f is non-zero function, and so

$$\lim_{\lambda \rightarrow \infty} q(\lambda) = \frac{c}{\lim_{\lambda \rightarrow \infty} p(\lambda)f(\lambda)} = \frac{c}{\text{constant}}.$$

Thus $\lim_{\lambda \rightarrow \infty} p(\lambda)f(\lambda)$ cannot be a zero.

□

- 4.8. A deterministic program is one that uses no random choices. Suppose \mathcal{L}_1 and \mathcal{L}_2 are two deterministic libraries with a common interface. Show that either $\mathcal{L}_1 \equiv \mathcal{L}_2$, or else \mathcal{L}_1 & \mathcal{L}_2 can be distinguished with advantage 1.

Solution. Since both \mathcal{L}_1 and \mathcal{L}_2 are deterministic libraries, they will always produce the same output for the same input, i.e., either

$$\mathcal{L}_1(x) = \mathcal{L}_2(x) \quad \text{or} \quad \mathcal{L}_1(x) \neq \mathcal{L}_2(x)$$

for any input x .

(i) ($\mathcal{L}_1(x) = \mathcal{L}_2(x)$) Clearly,

$$(\forall \text{input } x : \mathcal{L}_1(x) = \mathcal{L}_2(x)) \implies (\mathcal{L}_1 \equiv \mathcal{L}_2).$$

(ii) ($\mathcal{L}_1(x) \neq \mathcal{L}_2(x)$) Suppose that

$$\exists \text{input } x : \mathcal{L}_1(x) \neq \mathcal{L}_2(x).$$

We construct an adversary \mathcal{A} as follows:

- (a) $|\Pr[\mathcal{A} \diamond \mathcal{L}_1 \Rightarrow 1] - \Pr[\mathcal{A} \diamond \mathcal{L}_2 \Rightarrow 1]| = |1 - 0| = 1.$
- (b) $|\Pr[\mathcal{A} \diamond \mathcal{L}_1 \Rightarrow 1] - \Pr[\mathcal{A} \diamond \mathcal{L}_2 \Rightarrow 1]| = |0 - 1| = 1.$

□

4.12. Suppose you want to enforce password rules so that at least 2^{128} passwords satisfy the rules. How many characters long must the passwords be, in each of these cases?

- (a) Passwords consist of lowercase **a** through **z** only.
- (b) Passwords consist of lowercase and uppercase letters **a-z** and **A-Z**.
- (c) Passwords consist of lower/uppercase letters and digits **0-9**.
- (d) Passwords consist of lower/uppercase letters, digits, and any symbol characters that appear on a standard US keyboard (including the space character).

~	!	@	#	\$	%	^	&	*	()	_	+	Backspace
Tab	Q	W	E	R	T	Y	U	I	O	P	{	}	\
Caps Lock	A	S	D	F	G	H	J	K	L	:	"	'	Enter
Shift	Z	X	C	V	B	N	M	<	>	?	/	Shift	
Ctrl	Win	Alt								Alt	Win	Menu	Ctrl

Figure 3.2: Standard US Keyboard (<https://kbd-intl.narod.ru/english/layouts>)

Solution. We want to create a password system that allows for at least 2^{128} (16 bytes) different passwords.

- (a) We are only using lowercase letters **a-z**, which gives us **26** different possibilities for each character in the password. We need to solve the following equation for n (the length of the password):

$$26^n \geq 2^{128}.$$

Then

$$n \log(26) \geq 128 \log(2) \implies n \geq \frac{128 \cdot \log(2)}{\log(26)} \approx 27.2.$$

- (b)

$$52^n \geq 2^{128} \implies n \geq \frac{128 \cdot \log(2)}{\log(52)} \approx 22.4.$$

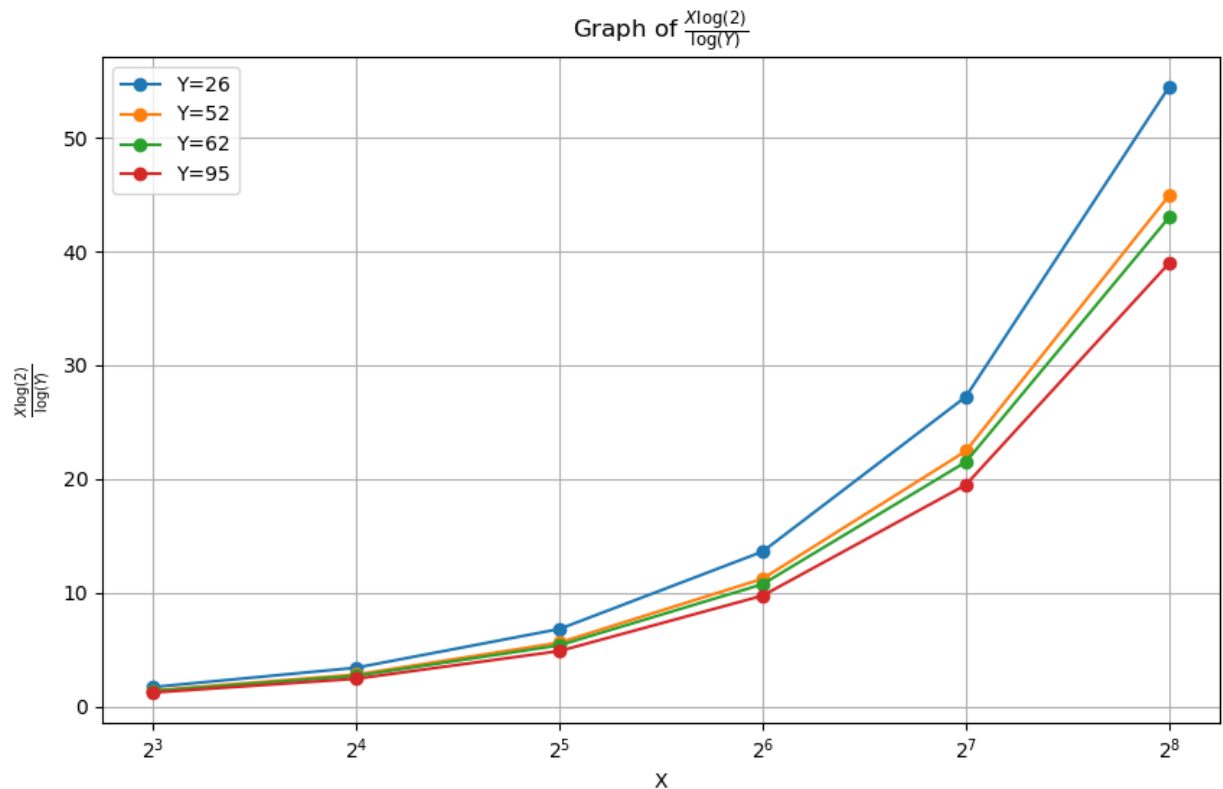
- (c)

$$62^n \geq 2^{128} \implies n \geq \frac{128 \cdot \log(2)}{\log(62)} \approx 21.5$$

- (d)

$$95^n \geq 2^{128} \implies n \geq \frac{128 \cdot \log(2)}{\log(95)} \approx 19.5$$

□



```

1 import matplotlib.pyplot as plt
2
3 # Given values of X and Y
4 X_values = [8, 16, 32, 64, 128, 256]
5 Y_values = [26, 52, 62, 95]
6
7 # Initialize a plot
8 plt.figure(figsize=(10,6))
9
10 # Loop through each Y value
11 for Y in Y_values:
12     # Calculate the expression for each X value
13     Z = [x * log(2) / log(Y) for x in X_values]
14
15     # Plot the result
16     plt.plot(X_values, Z, label='Y=' + str(Y), marker='o')
17
18 # Labeling the plot
19 plt.title(r'Graph of  $\frac{X \log(2)}{\log(Y)}$ ')
20 plt.xlabel('X')
21 plt.ylabel(r' $\frac{X \log(2)}{\log(Y)}$ ')
22 plt.xscale("log", base=2) # for logarithmic scale on x-axis
23 plt.legend()
24 plt.grid(True)
25 plt.show()

```

Chapter 4

Pseudo-random Generators (PRG)

4.1 Definition

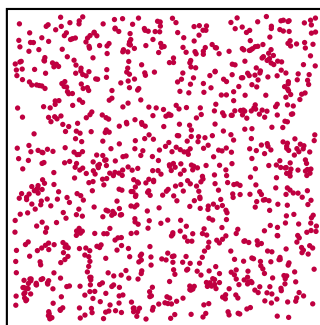
Pseudorandom Generator (PRG)

Definition 4.1. A deterministic function $G : \{0, 1\}^\lambda \rightarrow \{0, 1\}^{\lambda+l}$ with $l > 0$ is a **secure pseudorandom generator (PRG)** if $\mathcal{L}_{\text{PRG-real}}^G \approx \mathcal{L}_{\text{PRG-rand}}^G$, where:

$\mathcal{L}_{\text{PRG-real}}^G$	$\mathcal{L}_{\text{PRG-rand}}^G$
Query(): $s \leftarrow \{0, 1\}^\lambda$ return $G(s)$	Query(): $r \leftarrow \{0, 1\}^{\lambda+l}$ return r

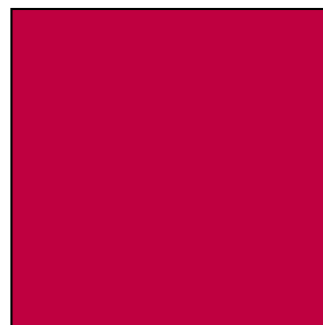
Remark 4.1. The value l is called the **stretch** of the PRG. The input s to the PRG is called a **seed**.

Remark 4.2. We illustrate the distributions, for a **length doubling** ($l = \lambda$) PRG (not drawn to scale):



$$\{0, 1\}^\lambda \rightarrow \{0, 1\}^{2\lambda}$$

Pseudorandom dist.



$$\{0, 1\}^{2\lambda}$$

Uniform dist.

Example 4.1 (Length-Doubling PRG). A straightforward approach for the PRG might be to duplicate its input string.

$$\boxed{\begin{array}{l} G(s): \\ \text{return } s \parallel s \end{array}}$$

For example, the following strings look likely they were sampled uniformly from $\{0, 1\}^8$:

11011101, 01110111, 01000100, ...

We can formalize this observation as an attack against the PRG-security of G :

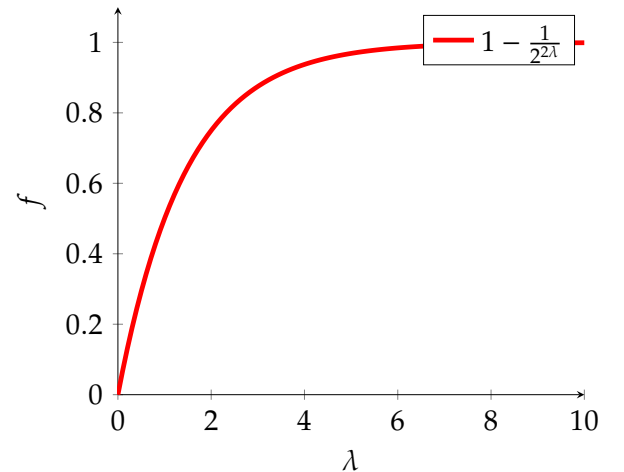
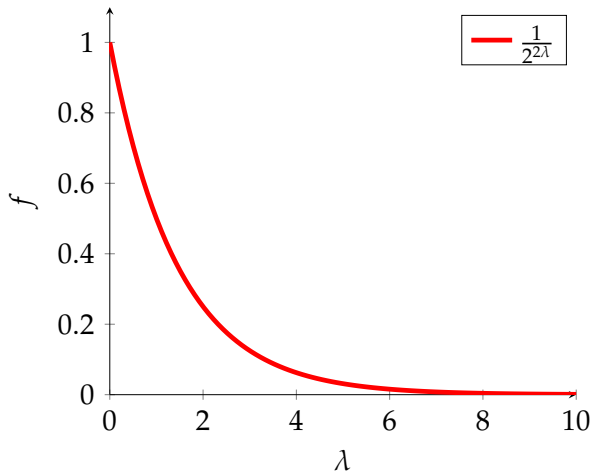
$$\boxed{\begin{array}{c} \mathcal{A} \\ x \parallel y := \text{Query}() \\ \text{return } x \stackrel{?}{=} y \end{array}}$$

Then Thus,

$$\begin{array}{c} \boxed{\begin{array}{c} \mathcal{A} \\ x \parallel y := \text{Query}() \\ \text{return } x \stackrel{?}{=} y \end{array}} \diamond \boxed{\begin{array}{c} \mathcal{L}_{\text{PRG-real}}^G \\ \text{Query}(): \\ s \leftarrow \{0, 1\}^\lambda \\ \text{return } G(s) \end{array}} \parallel \Pr[\mathcal{A} \diamond \mathcal{L}_{\text{PRG-real}}^G \Rightarrow 1] = 1. \\ \\ \boxed{\begin{array}{c} \mathcal{A} \\ x \parallel y := \text{Query}() \\ \text{return } x \stackrel{?}{=} y \end{array}} \diamond \boxed{\begin{array}{c} \mathcal{L}_{\text{PRG-rand}}^G \\ \text{Query}(): \\ r \leftarrow \{0, 1\}^{2\lambda} \\ \text{return } r \end{array}} \parallel \Pr[\mathcal{A} \diamond \mathcal{L}_{\text{PRG-rand}}^G \Rightarrow 1] = \frac{1}{2^{2\lambda}}. \end{array}$$

$$\text{Adv}_{\mathcal{A}} = \left| \Pr[\mathcal{A} \diamond \mathcal{L}_{\text{PRG-real}}^G \Rightarrow 1] - \Pr[\mathcal{A} \diamond \mathcal{L}_{\text{PRG-rand}}^G \Rightarrow 1] \right| = 1 - \frac{1}{2^{2\lambda}}$$

is non-negligible.



Comparison of Normalized Distributions:

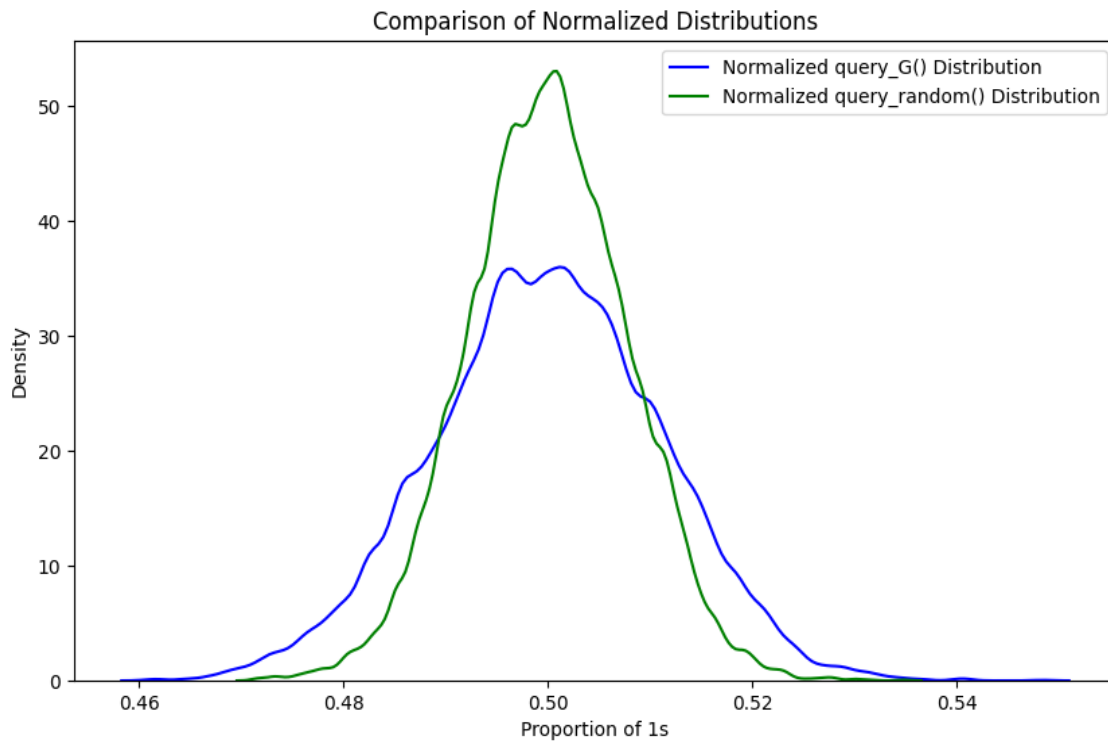


Figure 4.1: $\lambda = 1024$, i.e., $\{0, 1\}^{2048}$ with 100,000 experiments

```

1 import numpy as np
2 import matplotlib.pyplot as plt
3 import seaborn as sns
4
5 def G(s):
6     """
7     This function takes a list of bits (s), and returns a new list where each 8-
8     bit unit is doubled.
9     """
10    return s * 2
11
12 def query_G():
13     lambda_length = 32
14     s = np.random.randint(0, 2, lambda_length).tolist()
15     return G(s)
16
17 def query_random():
18     lambda_length = 1024
19     l_length = 1024
20     # Generates a list of 0s and 1s
21     r = np.random.randint(0, 2, lambda_length + l_length).tolist()
22     return r
23
24 # Define the number of experiments to run
25 num_experiments = 100000

```

```
26 # Record the outputs
27 outputs_G = [query_G() for _ in range(num_experiments)]
28 outputs_random = [query_random() for _ in range(num_experiments)]
29
30 # Convert outputs to the sum of their elements to see the distribution of the
    number of 1s
31 sums_G = [sum(output) for output in outputs_G]
32 sums_random = [sum(output) for output in outputs_random]
33
34 # Normalizing the sums by the length of the binary string
35 norm_sums_G = [s / 2048 for s in sums_G]
36 norm_sums_random = [s / 2048 for s in sums_random]
37
38 # Generate a Kernel Density Estimate plot for each normalized distribution
39 plt.figure(figsize=(10, 6))
40
41 # Plot KDE for normalized sums_G
42 sns.kdeplot(norm_sums_G, bw_adjust=0.5, label='Normalized query_G() Distribution', color='blue')
43
44 # Plot KDE for normalized sums_random
45 sns.kdeplot(norm_sums_random, bw_adjust=0.5, label='Normalized query_random() Distribution', color='green')
46
47 # Add a legend and titles
48 plt.legend()
49 plt.title('Comparison of Normalized Distributions')
50 plt.xlabel('Proportion of 1s')
51 plt.ylabel('Density')
52
53 plt.show()
```

4.2 Shorter Keys in One-Time-Secret Encryption

One-time Pad (OTP)

Construction 4.1. The **one-time pad** are given below:

$\mathcal{K} = \{0, 1\}^\lambda$	<u>KeyGen :</u>	<u>Enc($k, m \in \{0, 1\}^\lambda$) :</u>	<u>Dec($k, c \in \{0, 1\}^\lambda$) :</u>
$\mathcal{M} = \{0, 1\}^\lambda$	$k \xleftarrow{\$} \mathcal{K}$	return $k \oplus m$	return $k \oplus c$
$\mathcal{C} = \{0, 1\}^\lambda$	return k		

Pseudo-OTP

Construction 4.2. Let $G : \{0, 1\}^\lambda \rightarrow \{0, 1\}^{\lambda+l}$ be a PRG, and define the following:

$\mathcal{K} = \{0, 1\}^\lambda$	<u>KeyGen :</u>	<u>Enc(k, m) :</u>	<u>Dec(k, c) :</u>
$\mathcal{M} = \{0, 1\}^{\lambda+l}$	$k \xleftarrow{\$} \mathcal{K}$	return $G(k) \oplus m$	return $G(k) \oplus c$
$\mathcal{C} = \{0, 1\}^{\lambda+l}$	return k		

Computational One-Time Secrecy

Definition 4.2. An encryption scheme Σ has **(computational) one-time secrecy** if $\mathcal{L}_{\text{ots-1}}^\Sigma \approx \mathcal{L}_{\text{ots-2}}^\Sigma$. That is, if for all polynomial-time distinguishers \mathcal{A} , we have

$$\Pr \left[\mathcal{A} \diamond \mathcal{L}_{\text{ots-1}}^\Sigma \Rightarrow 1 \right] \approx \Pr \left[\mathcal{A} \diamond \mathcal{L}_{\text{ots-2}}^\Sigma \Rightarrow 1 \right].$$

Remark 4.3. Σ has **one-time secrecy** if

$\mathcal{L}_{\text{ots-1}}^\Sigma$		$\mathcal{L}_{\text{ots-2}}^\Sigma$
<u>Eve($m_1, m_2 \in \Sigma.\mathcal{M}$):</u>	\equiv	<u>Eve($m_1, m_2 \in \Sigma.\mathcal{M}$):</u>
$k \leftarrow \Sigma.\text{KeyGen}$		$k \leftarrow \Sigma.\text{KeyGen}$
$c \leftarrow \Sigma.\text{Enc}(k, m_1)$		$c \leftarrow \Sigma.\text{Enc}(k, m_2)$
return c		return c

Theorem 4.1. Let pOTP denote [Construction 4.2](#). If one constructs the pOTP utilizing a secure pseudorandom generator G , then pOTP has computational one-time secrecy.

Proof. We must show that

$$\mathcal{L}_{\text{ots-1}}^{\text{pOTP}} \approx \mathcal{L}_{\text{ots-2}}^{\text{pOTP}}.$$

We will show that a sequence of hybrid libraries satisfying the following:

$$\mathcal{L}_{\text{ots-1}}^{\text{pOTP}} \equiv \mathcal{L}_{\text{hyp-1}} \approx \mathcal{L}_{\text{hyp-2}} \equiv \mathcal{L}_{\text{hyp-3}} \equiv \mathcal{L}_{\text{hyp-4}} \equiv \mathcal{L}_{\text{hyp-5}} \approx \mathcal{L}_{\text{hyp-6}} \equiv \mathcal{L}_{\text{ots-2}}^{\text{pOTP}}.$$

$\mathcal{L}_{\text{ots-1}}^{\text{pOTP}}$:

$\mathcal{L}_{\text{ots-1}}^{\text{pOTP}}$
$\text{Eve}(m_1, m_2 \in \{0, 1\}^{\lambda+l})$: $k \leftarrow \{0, 1\}^\lambda$ $c := G(k) \oplus m_1$ return c

$\mathcal{L}_{\text{hyp-1}}$:

$\text{Eve}(m_1, m_2)$: $z \leftarrow \text{Query}()$ $c := z \oplus m_1$ return c	\diamond <table border="1"> <thead> <tr> <th>$\mathcal{L}_{\text{PRG-real}}^G$</th> </tr> </thead> <tbody> <tr> <td> $\text{Query}()$: $s \leftarrow \{0, 1\}^\lambda$ return $G(s)$ </td> </tr> </tbody> </table>	$\mathcal{L}_{\text{PRG-real}}^G$	$\text{Query}()$: $s \leftarrow \{0, 1\}^\lambda$ return $G(s)$
$\mathcal{L}_{\text{PRG-real}}^G$			
$\text{Query}()$: $s \leftarrow \{0, 1\}^\lambda$ return $G(s)$			

$\mathcal{L}_{\text{hyp-2}}$:

$\text{Eve}(m_1, m_2)$: $z \leftarrow \text{Query}()$ $c := z \oplus m_1$ return c	\diamond <table border="1"> <thead> <tr> <th>$\mathcal{L}_{\text{PRG-rand}}^G$</th> </tr> </thead> <tbody> <tr> <td> $\text{Query}()$: $r \leftarrow \{0, 1\}^{\lambda+l}$ return r </td> </tr> </tbody> </table>	$\mathcal{L}_{\text{PRG-rand}}^G$	$\text{Query}()$: $r \leftarrow \{0, 1\}^{\lambda+l}$ return r
$\mathcal{L}_{\text{PRG-rand}}^G$			
$\text{Query}()$: $r \leftarrow \{0, 1\}^{\lambda+l}$ return r			

$\mathcal{L}_{\text{hyp-3}}$:

$\mathcal{L}_{\text{ots-1}}^{\text{OTP}}$
$\text{Eve}(m_1, m_2)$: $z \leftarrow \{0, 1\}^{\lambda+l}$ $c := z \oplus m_1$ return c

$\mathcal{L}_{\text{ots-2}}^{\text{pOTP}}$:

$\mathcal{L}_{\text{ots-2}}^{\text{pOTP}}$
$\text{Eve}(m_1, m_2 \in \{0, 1\}^{\lambda+l})$: $k \leftarrow \{0, 1\}^\lambda$ $c := G(k) \oplus m_2$ return c

$\mathcal{L}_{\text{hyp-6}}$:

$\text{Eve}(m_1, m_2)$: $z \leftarrow \text{Query}()$ $c := z \oplus m_2$ return c	\diamond <table border="1"> <thead> <tr> <th>$\mathcal{L}_{\text{PRG-real}}^G$</th> </tr> </thead> <tbody> <tr> <td> $\text{Query}()$: $s \leftarrow \{0, 1\}^\lambda$ return $G(s)$ </td> </tr> </tbody> </table>	$\mathcal{L}_{\text{PRG-real}}^G$	$\text{Query}()$: $s \leftarrow \{0, 1\}^\lambda$ return $G(s)$
$\mathcal{L}_{\text{PRG-real}}^G$			
$\text{Query}()$: $s \leftarrow \{0, 1\}^\lambda$ return $G(s)$			

$\mathcal{L}_{\text{hyp-5}}$:

$\text{Eve}(m_1, m_2)$: $z \leftarrow \text{Query}()$ $c := z \oplus m_2$ return c	\diamond <table border="1"> <thead> <tr> <th>$\mathcal{L}_{\text{PRG-rand}}^G$</th> </tr> </thead> <tbody> <tr> <td> $\text{Query}()$: $r \leftarrow \{0, 1\}^{\lambda+l}$ return r </td> </tr> </tbody> </table>	$\mathcal{L}_{\text{PRG-rand}}^G$	$\text{Query}()$: $r \leftarrow \{0, 1\}^{\lambda+l}$ return r
$\mathcal{L}_{\text{PRG-rand}}^G$			
$\text{Query}()$: $r \leftarrow \{0, 1\}^{\lambda+l}$ return r			

$\mathcal{L}_{\text{hyp-4}}$:

$\mathcal{L}_{\text{ots-2}}^{\text{OTP}}$
$\text{Eve}(m_1, m_2)$: $z \leftarrow \{0, 1\}^{\lambda+l}$ $c := G(k) \oplus m_2$ return c

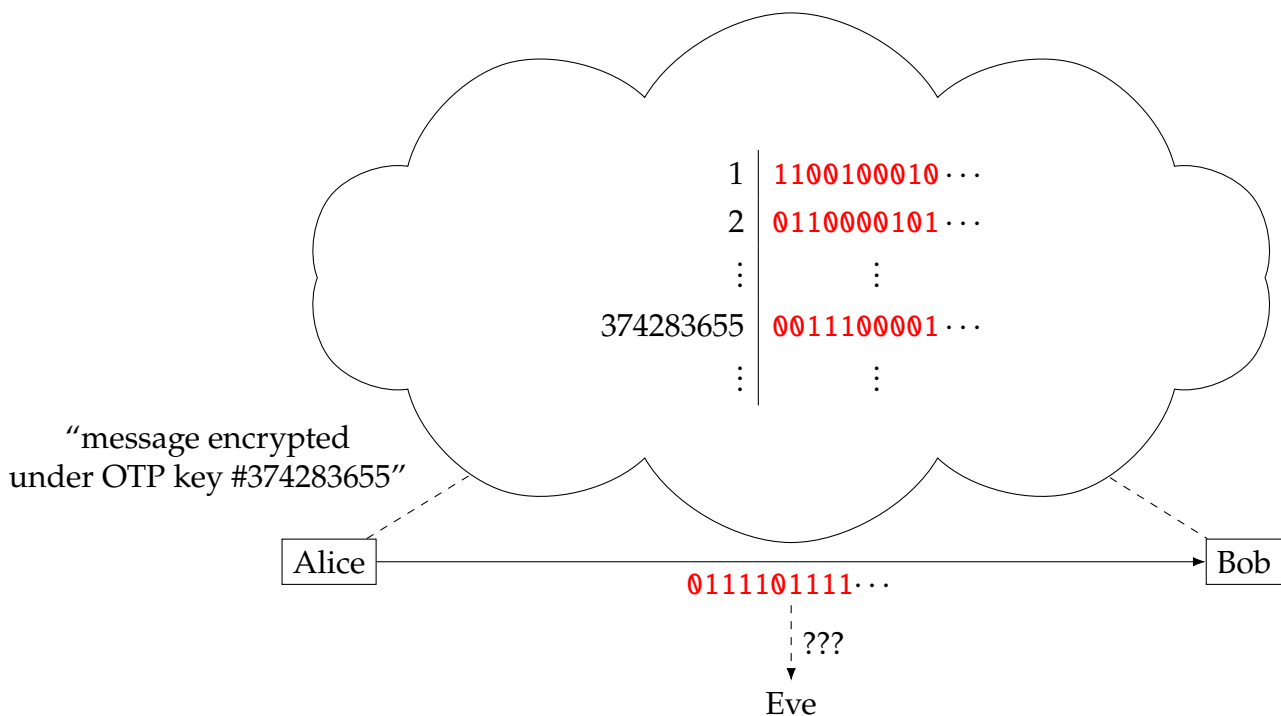
Chapter 5

Pseudo-Random Functions & Block Ciphers

Alice then informs Bob that chunk n has been used, without revealing the actual bits of r_n . Bob, having access to the shared randomness \mathcal{R} , can decrypt the message by computing:

$$m = c \oplus r_n$$

Since the eavesdropper does not possess \mathcal{R} , and given that each r_n is used only once, the encrypted message c reveals no information about the message m as long as the XOR operation with r_n is uniformly distributed. Thus, the encryption scheme is information-theoretically secure.

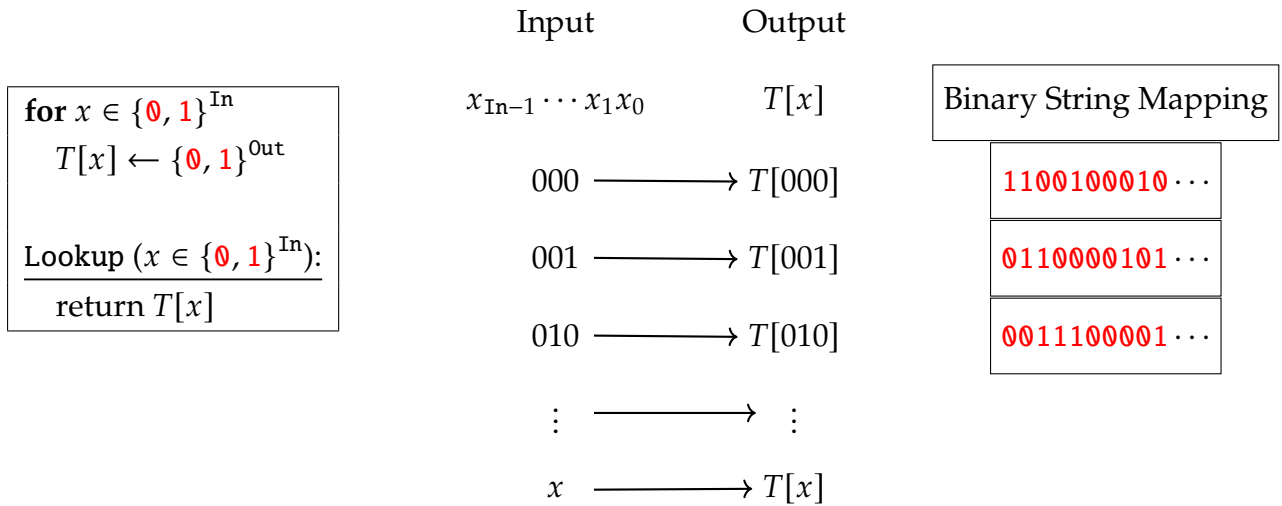


While the notion of infinite shared randomness is impractical, it can be approximated by an exponential amount of shared resources. Consider a table \mathcal{T} shared between Alice and Bob, containing 2^λ unique one-time pad keys, sufficient for an extensive number of message encryptions.

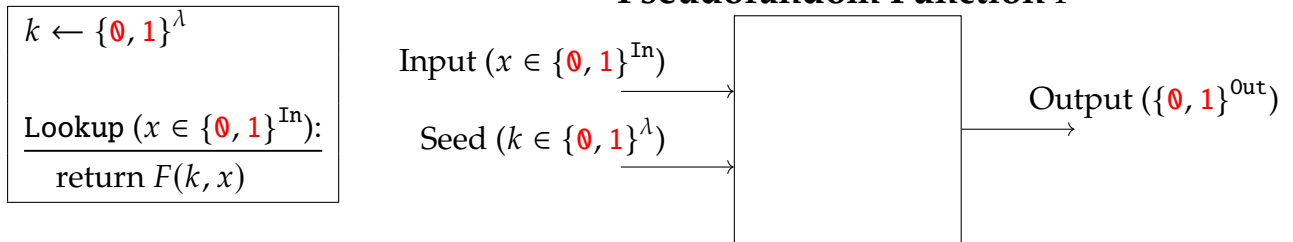
5.1 Definition

The goal of a pseudorandom function is to “look like” a uniformly chosen array / lookup table.

Array T



Pseudorandom Function F



A Pseudo-Random Function (PRF) is a fundamental concept in cryptography, typically defined in the context of a family of functions. Let’s denote a PRF family by F , where each function f_k in F is indexed by a key k from a key space K . The function f_k maps inputs from an input space X to outputs in an output space Y . Mathematically, for a key k , the PRF is defined as:

$$f_k : X \rightarrow Y$$

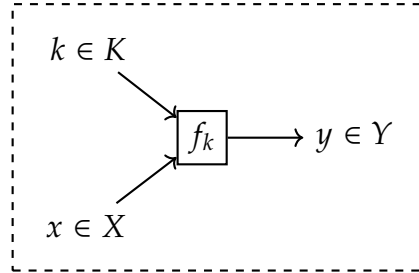
Given k , for any input $x \in X$, $f_k(x)$ is easy to compute. However, without knowledge of k , the function’s output is indistinguishable from a truly random function from an adversary’s point of view, given polynomially bounded computational resources.

A truly random function (RF), on the other hand, is a function where every possible input $x \in X$ is mapped to an output $y \in Y$ completely at random, without any deterministic process. Formally, an RF is defined as a function $h : X \rightarrow Y$ where each $h(x)$ is chosen uniformly at random from Y .

The key distinction between a PRF and an RF is that a PRF’s output is reproducible given the same key and input, whereas an RF provides no such guarantee—the output for the same input can vary with each function invocation.

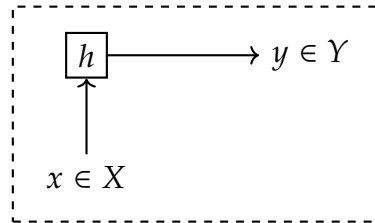
Remark 5.1.

Pseudo-Random Function (PRF)



Output is deterministic with the same key and input

Random Function (RF)



Output varies randomly for the same input

PRF Security

Definition 5.1. A deterministic function

$$F : \{0, 1\}^\lambda \times \{0, 1\}^{\text{In}} \rightarrow \{0, 1\}^{\text{Out}}$$

is a **secure pseudo-random function (PRF)** if $\mathcal{L}_{\text{PRF-real}}^F \approx \mathcal{L}_{\text{PRF-rand}}^F$, where

$\mathcal{L}_{\text{PRF-real}}^F$
$k \leftarrow \{0, 1\}^\lambda$
<u>Lookup ($x \in \{0, 1\}^{\text{In}}$):</u>
return $F(k, x)$

$\mathcal{L}_{\text{PRF-rand}}^F$
$T := \{\}$
<u>Lookup ($x \in \{0, 1\}^{\text{In}}$):</u>
if $T[x]$ undefined:
$T[x] \leftarrow \{0, 1\}^{\text{Out}}$
return $T[x]$

Algorithm 8: $\mathcal{L}_{\text{PRF-real}}^F$ and $\mathcal{L}_{\text{PRF-rand}}^F$

```

1: Function  $\mathcal{L}_{\text{PRF-real}}^F$ :
2:    $k \leftarrow \{0, 1\}^\lambda$ 
3:   Function Lookup( $x \in \{0, 1\}^n$ ):
4:     return  $F(k, x)$ 
5:   end
6: end
7: Function  $\mathcal{L}_{\text{PRF-rand}}^F$ :
8:    $T := \{\}$ 
9:   Function Lookup( $x \in \{0, 1\}^n$ ):
10:    if  $T[x]$  undefined then
11:       $T[x] \leftarrow \{0, 1\}^{\text{out}}$ 
12:    end
13:    return  $T[x]$ 
14:  end
15: end

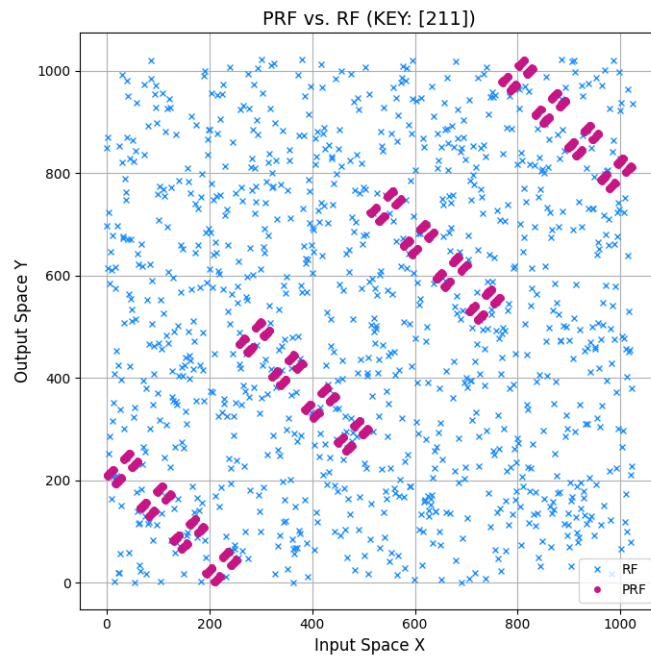
```

Example 5.1 (How NOT to Build a PRF). Suppose we have a length-doubling PRG

$$G : \{0, 1\}^\lambda \rightarrow \{0, 1\}^{2\lambda}$$

and try to use it to construct a PRF F as follows:

$F(k, x)$:
 return $G(k) \oplus x$



```

1 import matplotlib.pyplot as plt
2 import numpy as np
3
4 # Define a class RandomFunction to simulate a random function.
5 class RandomFunction:
6     def __init__(self, y_space): # Initialize with a range of possible y values
7         self.y_space = y_space
8         self.results = {} # Dictionary to store x:y mappings
9
10    def __call__(self, x): # If x is new, assign a random y from y_space, else
        return stored y
11        if x not in self.results:
12            self.results[x] = np.random.choice(self.y_space)
13        return self.results[x]
14
15 # Generate a random key for the pseudorandom function (PRF)
16 key = np.random.randint(0, 2**10, size=1)
17
18 def prf(x, k): # Define a simple PRF: bitwise XOR of input x with key k
19     x_k = np.bitwise_xor(x, k)
20     return x_k
21
22 # Create an input space for the functions
23 input_space = np.arange(2**10)
24
25 # Calculate PRF output for each input value
26 prf_output = np.array([prf(x, key) for x in input_space])
27
28 # Define the output space for the random function
29 output_space = np.arange(2**10)
30
31 # Create an instance of RandomFunction with the defined output space
32 rf_instance = RandomFunction(output_space)
33
34 # Calculate RF output for each input value
35 rf_output = np.array([rf_instance(x) for x in input_space])
36
37 # Set up the plot with a specific size
38 plt.figure(figsize=(7, 7))
39
40 # Plot RF output vs input space
41 plt.plot(input_space, rf_output, 'x', color='dodgerblue', label='RF', markersize
         =4)
42 # Plot PRF output vs input space
43 plt.plot(input_space, prf_output, 'o', color='mediumvioletred', label='PRF',
         markersize=4)
44
45 plt.title(f'PRF vs. RF (KEY: {key})', fontsize=14)
46 plt.xlabel('Input Space X', fontsize=12)
47 plt.ylabel('Output Space Y', fontsize=12)
48 plt.legend()
49
50 plt.grid(True)
51
52 plt.tight_layout()
53 plt.show()

```

Chapter 6

Security Against Chosen Plaintext Attacks

Our earlier rules for secure encryption were about using a key for just one message. But it's better to have a method where you can encrypt many messages with the same key.

We just make the libraries pick a secret key once, and this key is used for all messages. To explain this in a more detailed way:

CPA security

Definition 6.1. Let Σ be an encryption scheme. We say that Σ has **CPA security** (security against chosen-plaintext attacks) if $\mathcal{L}_{\text{cpa-L}}^{\Sigma} \approx \mathcal{L}_{\text{cpa-R}}^{\Sigma}$.

Algorithm 9: $\mathcal{L}_{\text{cpa-L}}^{\Sigma}$

Input: $m_L, m_R \in \Sigma.M$

Output: Ciphertext c

```
1  $k \leftarrow \Sigma.\text{KeyGen}$ 
2 Function EAVESDROP( $m_L, m_R \in \Sigma.M$ ):
3    $c \leftarrow \Sigma.\text{Enc}(k, m_L)$ 
4   return  $c$ 
5 end
```

Algorithm 10: $\mathcal{L}_{\text{cpa-R}}^{\Sigma}$

Input: $m_L, m_R \in \Sigma.M$

Output: Ciphertext c

```
1  $k \leftarrow \Sigma.\text{KeyGen}$ 
2 Function EAVESDROP( $m_L, m_R \in \Sigma.M$ ):
3    $c \leftarrow \Sigma.\text{Enc}(k, m_R)$ 
4   return  $c$ 
5 end
```

Note. CPA security is often called “IND-CPA” security, meaning “indistinguishability of ciphertexts under chosen-plaintext attack.”

6.1 Limits of Deterministic Encryption

We already know about block ciphers, also called PRPs. These look like they have everything necessary for safe encryption. In a block cipher F is used for encrypting, F^{-1} is used for decrypting, and the results of F appear to be random, but they're not truly random. What else would you need in a reliable encryption method?

Example 6.1. We will see that a block cipher, when used "as-is," is not a CPA-secure encryption scheme. Let F denote the block cipher and suppose its block length is blen .

Consider the following adversary \mathcal{A} , that tries to distinguish the \mathcal{L} libraries:

Algorithm 11: Adversary \mathcal{A} for $\mathcal{L}_{\text{CPA-}^*}$

Input: Two arguments $\mathbf{0}^{\text{blen}}, \mathbf{1}^{\text{blen}}$

Output: True or False

```

1  $c_1 \leftarrow \text{EAVESDROP}(\mathbf{0}^{\text{blen}}, \mathbf{0}^{\text{blen}});$ 
2  $c_2 \leftarrow \text{EAVESDROP}(\mathbf{0}^{\text{blen}}, \mathbf{1}^{\text{blen}});$ 
3 return  $c_1 = c_2;$ 

```

Algorithm 12: EAVESDROP Left and Right

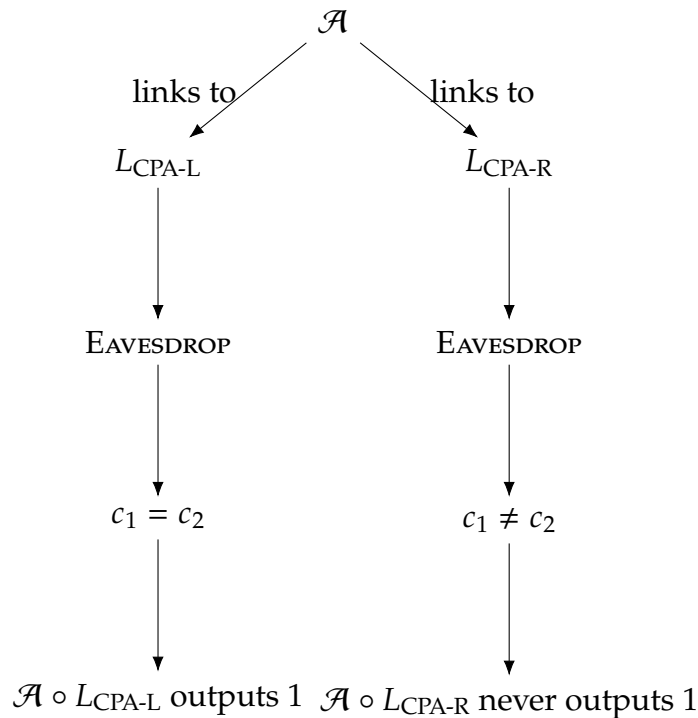
Input: $k \leftarrow \{\mathbf{0}, \mathbf{1}\}^\lambda, m_1, m_2$

Output: Cipher c

```

1 EAVESDROP Left:
2  $c \leftarrow F(k, m_1);$ 
3 return  $c;$ 
4 EAVESDROP Right:
5  $c \leftarrow F(k, m_2);$ 
6 return  $c;$ 

```



When \mathcal{A} is linked to $\mathcal{L}_{\text{CPA-L}}$, the EAVESDROP algorithm will encrypt its first argument. So, c_1 and c_2 will both be computed as $F(k, \mathbf{0}^{\text{blen}})$. Since F is a deterministic function, this results in identical outputs from EAVESDROP. In other words $c_1 = c_2$, and $\mathcal{A} \circ \mathcal{L}_{\text{CPA-L}}$ always outputs 1.

When \mathcal{A} is linked to $\mathcal{L}_{\text{CPA-R}}$, the EAVESDROP algorithm will encrypt its second argument. So, c_1 and c_2 are computed as $c_1 = F(k, \mathbf{0}^{\text{blen}})$ and $c_2 = F(k, \mathbf{1}^{\text{blen}})$. Since F is a permutation, $c_1 \neq c_2$, so $\mathcal{A} \circ \mathcal{L}_{\text{CPA-R}}$ never outputs 1.

Bibliography

- [1] M. Rosulek, *The Joy of Cryptography*, [Online]. Available: <https://joyofcryptography.com>
- [2] N. P. Smart, *Cryptography Made Simple*. 1st ed. Springer International Publishing, 2016.
- [3] J. Katz and Y. Lindell, *Introduction to Modern Cryptography*. 2nd ed. Chapman and Hall/CRC, 2014.