Chapter 2

Perfectly Secret Encryption

In the previous chapter we presented historical encryption schemes and showed how they can be completely broken with very little computational effort. In this chapter, we look at the other extreme and study encryption schemes that are *provably secure* even against an adversary who has unbounded computational power. Such schemes are called *perfectly secret*. Besides rigorously defining the notion, we will explore conditions under which perfect secrecy can and cannot be achieved.

The material in this chapter belongs, in some sense, more to the world of "classical" cryptography than to the world of "modern" cryptography. Besides the fact that all the material introduced here was developed before the revolution in cryptography that took place in the mid-'70s and '80s, the constructions we study in this chapter rely only on the first and third principles outlined in Section 1.4. That is, precise mathematical definitions are used and rigorous proofs are given, but it will not be necessary to rely on any unproven assumptions. It is clearly advantageous to avoid assumptions; we will see, however, that doing so has inherent limitations. Thus, in addition to serving as a good basis for understanding the principles underlying modern cryptography, the results of this chapter also justify our later adoption of all three of the aforementioned principles.

Beginning with this chapter, we assume familiarity with basic probability theory. The relevant notions are reviewed in Appendix A.3.

2.1 Definitions

We begin by recalling and expanding upon the syntax that was introduced in the previous chapter. An encryption scheme is defined by three algorithms Gen, Enc, and Dec, as well as a specification of a (finite) *message space* $\mathcal{M}$ with $|\mathcal{M}| > 1$.[1] The key-generation algorithm Gen is a probabilistic algorithm that outputs a key k chosen according to some distribution. We denote by $\mathcal{K}$ the (finite) *key space*, i.e., the set of all possible keys that can be output by Gen. The encryption algorithm Enc takes as input a key $k \in \mathcal{K}$ and a

[1] If $|\mathcal{M}| = 1$ there is only one message and no point in communicating, let alone encrypting.

message $m \in \mathcal{M}$, and outputs a ciphertext c. The encryption algorithm may be probabilistic (so that $\mathsf{Enc}_k(m)$ might output a different ciphertext when run multiple times), and so we write $c \leftarrow \mathsf{Enc}_k(m)$ to denote the possibly probabilistic process by which message m is encrypted using key k to give ciphertext c. (In case Enc is deterministic, we may emphasize this by writing $c := \mathsf{Enc}_k(m)$. Looking ahead, we also use the notation $x \leftarrow S$ to denote uniform selection of x from a set S.) We let $\mathcal{C}$ denote the set of all possible ciphertexts that can be output by $\mathsf{Enc}_k(m)$, for all possible choices of $k \in \mathcal{K}$ and $m \in \mathcal{M}$ (and for all random choices of Enc in case it is randomized). The decryption algorithm Dec takes as input a key $k \in \mathcal{K}$ and a ciphertext $c \in \mathcal{C}$ and outputs a message $m \in \mathcal{M}$. We assume *perfect correctness*, meaning that for all $k \in \mathcal{K}$, $m \in \mathcal{M}$, and any ciphertext c output by $\mathsf{Enc}_k(m)$, it holds that $\mathsf{Dec}_k(c) = m$ with probability 1. Perfect correctness implies that we may assume Dec is deterministic without loss of generality, since $\mathsf{Dec}_k(c)$ must give the same output every time it is run. We will thus write $m := \mathsf{Dec}_k(c)$ to denote the process of decrypting ciphertext c using key k to yield the message m as the result.

In the definitions and theorems below, we refer to probability distributions over $\mathcal{K}$, $\mathcal{M}$, and $\mathcal{C}$. The distribution over $\mathcal{K}$ is the one defined by running Gen and taking the output. (It is almost always the case that Gen chooses a key uniformly from $\mathcal{K}$ and, in fact, we may assume this without loss of generality; see Exercise 2.1.) We let K be a random variable denoting the value of the key output by Gen; thus, for any $k \in \mathcal{K}$, $\Pr[K = k]$ denotes the probability that the key output by Gen is equal to k. Similarly, we let M be a random variable denoting the message being encrypted, so $\Pr[M = m]$ denotes the probability that the message M takes on the value $m \in \mathcal{M}$. The probability distribution of the message is not determined by the encryption scheme itself, but instead reflects the likelihood of different messages being sent by the parties using the scheme, as well as an adversary's uncertainty about what will be sent. As an example, an adversary may know that the message will either be `attack today` or `don't attack`. The adversary may even know (by other means) that with probability 0.7 the message will be a command to attack and with probability 0.3 the message will be a command not to attack. In this case, we have $\Pr[M = \texttt{attack today}] = 0.7$ and $\Pr[M = \texttt{don't attack}] = 0.3$.

K and M are assumed to be independent, i.e., what is being communicated by the parties is independent of the key they happen to share. This makes sense, among other reasons, because the distribution over $\mathcal{K}$ is determined by the encryption scheme itself (since it is defined by Gen), while the distribution over $\mathcal{M}$ depends on the context in which the encryption scheme is being used.

Fixing an encryption scheme and a distribution over $\mathcal{M}$ determines a distribution over the space of ciphertexts $\mathcal{C}$ given by choosing a key $k \in \mathcal{K}$ (according to Gen) and a message $m \in \mathcal{M}$ (according to the given distribution), and then computing the ciphertext $c \leftarrow \mathsf{Enc}_k(m)$. We let C be the random variable denoting the resulting ciphertext and so, for $c \in \mathcal{C}$, write $\Pr[C = c]$ to denote the probability that the ciphertext is equal to the fixed value c.

Example 2.1

We work through a simple example for the shift cipher (cf. Section 1.3). Here, by definition, we have $\mathcal{K} = \{0, \ldots, 25\}$ with $\Pr[K = k] = 1/26$ for each $k \in \mathcal{K}$.

Say we are given the following distribution over $\mathcal{M}$:

$$\Pr[M = \mathtt{a}] = 0.7 \quad \text{and} \quad \Pr[M = \mathtt{z}] = 0.3.$$

What is the probability that the ciphertext is B? There are only two ways this can occur: either $M = \mathtt{a}$ and $K = 1$, or $M = \mathtt{z}$ and $K = 2$. By independence of M and K, we have

$$\begin{aligned}\Pr[M = \mathtt{a} \wedge K = 1] &= \Pr[M = \mathtt{a}] \cdot \Pr[K = 1] \\ &= 0.7 \cdot \left(\frac{1}{26}\right).\end{aligned}$$

Similarly, $\Pr[M = \mathtt{z} \wedge K = 2] = 0.3 \cdot \left(\frac{1}{26}\right)$. Therefore,

$$\begin{aligned}\Pr[C = \mathtt{B}] &= \Pr[M = \mathtt{a} \wedge K = 1] + \Pr[M = \mathtt{z} \wedge K = 2] \\ &= 0.7 \cdot \left(\frac{1}{26}\right) + 0.3 \cdot \left(\frac{1}{26}\right) \; = \; 1/26.\end{aligned}$$

We can calculate conditional probabilities as well. For example, what is the probability that the message a was encrypted, given that we observe ciphertext B? Using Bayes' Theorem (Theorem A.8) we have

$$\begin{aligned}\Pr[M = \mathtt{a} \mid C = \mathtt{B}] &= \frac{\Pr[M = \mathtt{a} \wedge C = \mathtt{B}]}{\Pr[C = \mathtt{B}]} \\ &= \frac{\Pr[C = \mathtt{B} \mid M = \mathtt{a}] \cdot \Pr[M = \mathtt{a}]}{\Pr[C = \mathtt{B}]} \\ &= \frac{0.7 \cdot \Pr[C = \mathtt{B} \mid M = \mathtt{a}]}{1/26}.\end{aligned}$$

Note that $\Pr[C = \mathtt{B} \mid M = \mathtt{a}] = 1/26$, since if $M = \mathtt{a}$ then the only way $C = \mathtt{B}$ can occur is if $K = 1$ (which occurs with probability $1/26$). We conclude that $\Pr[M = \mathtt{a} \mid C = \mathtt{B}] = 0.7$. ◇

Example 2.2

Consider the shift cipher again, but with the following distribution over $\mathcal{M}$:

$$\Pr[M = \mathtt{kim}] = 0.5, \; \Pr[M = \mathtt{ann}] = 0.2, \; \Pr[M = \mathtt{boo}] = 0.3.$$

What is the probability that $C = \mathtt{DQQ}$? The only way this ciphertext can occur is if $M = \mathtt{ann}$ and $K = 3$, or $M = \mathtt{boo}$ and $K = 2$, which happens with probability $0.2 \cdot 1/26 + 0.3 \cdot 1/26 = 1/52$.

So what is the probability that ann was encrypted, conditioned on observing the ciphertext DQQ? A calculation as above using Bayes' Theorem gives $\Pr[M = \mathtt{ann} \mid C = \mathtt{DQQ}] = 0.2/0.5 = 0.4$. ◇

Perfect secrecy. We are now ready to define the notion of *perfect secrecy*. We imagine an adversary who knows the probability distribution over $\mathcal{M}$; that is, the adversary knows the likelihood that different messages will be sent. This adversary also knows the encryption scheme being used; the only thing unknown to the adversary is the key shared by the parties. A message is chosen by one of the honest parties and encrypted, and the resulting ciphertext transmitted to the other party. The adversary can *eavesdrop* on the parties' communication, and thus observe this ciphertext. (Using the terminology from the previous chapter, this is a one-time, ciphertext-only attack.) For a scheme to be perfectly secret, observing this ciphertext should have *no effect* on the adversary's knowledge regarding the actual message that was sent; in other words, the *a posteriori* probability that some message $m \in \mathcal{M}$ was sent, conditioned on the ciphertext that was observed, should be no different from the *a priori* probability that m would be sent. This means that the ciphertext reveals nothing about the underlying plaintext, and the adversary learns absolutely nothing about the plaintext that was encrypted. Formally:

DEFINITION 2.3 *An encryption scheme* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *with message space* $\mathcal{M}$ *is* perfectly secret *if for every probability distribution over* $\mathcal{M}$, *every message* $m \in \mathcal{M}$, *and every ciphertext* $c \in \mathcal{C}$ *for which* $\Pr[C = c] > 0$*:*

$$\Pr[M = m \mid C = c] = \Pr[M = m].$$

(The requirement that $\Pr[C = c] > 0$ is a technical one needed to prevent conditioning on a zero-probability event.)

Perfect indistinguishability. We now give an equivalent formulation of perfect secrecy. Informally, this formulation states that, for a perfectly secret encryption scheme, the probability distribution over $\mathcal{C}$ does not depend on the plaintext. I.e., for any two messages $m_0, m_1 \in \mathcal{M}$ the distribution of the ciphertext when m_0 is encrypted is identical to the distribution of the ciphertext when m_1 is encrypted. This is just another way of saying that the ciphertext contains no information about the plaintext. We refer to this formulation as *perfect indistinguishability* because it implies that it is impossible to distinguish an encryption of m_0 from an encryption of m_1 (since the distribution over the ciphertext is the same in each case).

LEMMA 2.4 *An encryption scheme* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *with message space* $\mathcal{M}$ *is perfectly secret if and only if for every* $m_0, m_1 \in \mathcal{M}$, *and every* $c \in \mathcal{C}$*:*

$$\Pr[\mathsf{Enc}_K(m_0) = c] = \Pr[\mathsf{Enc}_K(m_1) = c].$$

PROOF We show that if the stated condition holds, then the scheme is perfectly secret; the converse implication is left to Exercise 2.4. Fix a distribution over $\mathcal{M}$, a message m, and a ciphertext c for which $\Pr[C = c] > 0$. If $\Pr[M = m] = 0$ then we trivially have

$$\Pr[M = m \mid C = c] = 0 = \Pr[M = m].$$

So, assume $\Pr[M = m] > 0$. Notice first that

$$\Pr[C = c \mid M = m] = \Pr[\mathsf{Enc}_K(M) = c \mid M = m] = \Pr[\mathsf{Enc}_K(m) = c],$$

where the first equality is by definition of the random variable C, and the second is because we condition on the event that M is equal to m. If the condition of the lemma holds, then $\Pr[\mathsf{Enc}_K(m) = c]$ is independent of m and so we may define the constant $\delta_c \stackrel{\text{def}}{=} \Pr[\mathsf{Enc}_K(m) = c] = \Pr[C = c \mid M = m]$. Now, using Bayes' Theorem (see Appendix A.3) we have

$$\begin{aligned}
\Pr[M = m \mid C = c] &= \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\Pr[C = c]} \\
&= \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\sum_{m' \in \mathcal{M}} \Pr[C = c \mid M = m'] \cdot \Pr[M = m']} \\
&= \frac{\delta_c \cdot \Pr[M = m]}{\sum_{m' \in \mathcal{M}} \delta_c \cdot \Pr[M = m']} \\
&= \frac{\Pr[M = m]}{\sum_{m' \in \mathcal{M}} \Pr[M = m']} \;=\; \Pr[M = m],
\end{aligned}$$

where the sum in the denominator is over $m' \in \mathcal{M}$ with $\Pr[M = m'] \neq 0$. We thus see that the scheme is perfectly secret. ■

Adversarial indistinguishability. We conclude this section by presenting another equivalent definition of perfect secrecy. This definition is based on an *experiment* involving an adversary passively observing a ciphertext, and then trying to guess which of two possible messages was encrypted. We introduce this notion since it will serve as our starting point when we define computational security in the next chapter. Indeed, throughout the rest of the book we will often use experiments of this sort to define security.

In the present context, we consider the following experiment: an adversary $\mathcal{A}$ first chooses two arbitrary messages $m_0, m_1 \in \mathcal{M}$. One of these two messages is chosen uniformly at random and encrypted using a random key; the resulting ciphertext is given to $\mathcal{A}$. Finally, $\mathcal{A}$ outputs a "guess" as to which of the two messages was encrypted; $\mathcal{A}$ *succeeds* if it guesses correctly. An encryption scheme is *adversarially indistinguishable* if *no* adversary can succeed with probability better than 1/2. (Note that, regardless of the encryption scheme being used, $\mathcal{A}$ can succeed with probability 1/2 by outputting a random guess; the requirement is that no attacker can do better.)

Formally, let $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ be an encryption scheme defined for a message space $\mathcal{M}$. Let $\mathcal{A}$ be an adversary, which is formally just a (stateful) algorithm. We define an experiment $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi}$ as follows:

The adversarial indistinguishability experiment $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi}$:

1. *The adversary $\mathcal{A}$ outputs a pair of messages $m_0, m_1 \in \mathcal{M}$.*
2. *A key k is generated using Gen, and a uniform bit $b \leftarrow \{0,1\}$ is chosen. Ciphertext $c \leftarrow \mathsf{Enc}_k(m_b)$ is computed and given to $\mathcal{A}$. We refer to c as the* challenge ciphertext.
3. *$\mathcal{A}$ outputs a bit b'.*
4. *The output of the experiment is defined to be 1 if $b' = b$, and 0 otherwise. We write $\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1$ if the output of the experiment is 1 and in this case we say that $\mathcal{A}$* succeeds.

DEFINITION 2.5 *Encryption scheme $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ with message space $\mathcal{M}$ is* adversarially indistinguishable *if for every $\mathcal{A}$ it holds that*

$$\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] = \frac{1}{2}.$$

The following lemma states that Definition 2.5 is equivalent to Definition 2.3. We leave the proof of the lemma as Exercise 2.5.

LEMMA 2.6 *Let $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ be an encryption scheme with message space $\mathcal{M}$. Then Π is perfectly secret if and only if it is adversarially indistinguishable.*

Example 2.7

We show that the Vigenère cipher is not perfectly secret, at least for certain parameters. Concretely, let Π denote the Vigenère cipher for the message space of two-character strings, and where the period is chosen uniformly in $\{1,2\}$. To show that Π is not perfectly secret, we show an $\mathcal{A}$ for which $\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] > \frac{1}{2}$. Consider the following adversary $\mathcal{A}$:

1. Output $m_0 = \mathsf{aa}$ and $m_1 = \mathsf{ab}$.
2. Upon receiving the challenge ciphertext $c = c_1c_2$, do the following: if $c_1 = c_2$ output 0; else output 1.

Computation of $\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right]$ is tedious but straightforward.

$$\begin{aligned}
&\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] \\
&= \frac{1}{2} \cdot \Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1 \mid b = 0\right] + \frac{1}{2} \cdot \Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1 \mid b = 1\right] \\
&= \frac{1}{2} \cdot \Pr[\mathcal{A} \text{ outputs } 0 \mid b = 0] + \frac{1}{2} \cdot \Pr[\mathcal{A} \text{ outputs } 1 \mid b = 1], \qquad (2.1)
\end{aligned}$$

where b is the uniform bit determining which message gets encrypted. $\mathcal{A}$ outputs 0 if and only if the two characters of the ciphertext $c = c_1c_2$ are equal. When $b = 0$ (so $m_0 = \mathtt{aa}$ is encrypted) then $c_1 = c_2$ if either (1) a key of period 1 is chosen, or (2) a key of period 2 is chosen, and both characters of the key are equal. The former occurs with probability $\frac{1}{2}$, and the latter occurs with probability $\frac{1}{2} \cdot \frac{1}{26}$. So

$$\Pr[\mathcal{A} \text{ outputs } 0 \mid b = 0] = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{26} \approx 0.52.$$

When $b = 1$ then $c_1 = c_2$ only if a key of period 2 is chosen and the first character of the key is one more than the second character of the key. So

$$\Pr[\mathcal{A} \text{ outputs } 1 \mid b = 1] = 1 - \Pr[\mathcal{A} \text{ outputs } 0 \mid b = 1] = 1 - \frac{1}{2} \cdot \frac{1}{26} \approx 0.98.$$

Plugging into Equation (2.1) then gives

$$\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] = \frac{1}{2} \cdot (0.52 + 0.98) \approx 0.75 > \frac{1}{2},$$

and the scheme is not perfectly secret. ◇

2.2 The One-Time Pad

In 1917, Vernam patented a perfectly secret encryption scheme now called the *one-time pad*. At the time Vernam proposed the scheme, there was no proof that it was perfectly secret; in fact, there was not yet a notion of what perfect secrecy was. Approximately 25 years later, however, Shannon introduced the notion of perfect secrecy and demonstrated that the one-time pad achieves this level of security.

In describing the scheme we let $a \oplus b$ denote the *bitwise exclusive-or* (XOR) of two binary strings a and b (i.e., if $a = a_1 \cdots a_\ell$ and $b = b_1 \cdots b_\ell$ are ℓ-bit strings, then $a \oplus b$ is the ℓ-bit string given by $a_1 \oplus b_1 \cdots a_\ell \oplus b_\ell$). In the one-time pad encryption scheme the key is a random string of the same length as the message; the ciphertext is computed by simply XORing the key and the message. A formal definition is given as Construction 2.8. Before discussing security, we first verify correctness: for every key k and every message m it holds that $\mathsf{Dec}_k(\mathsf{Enc}_k(m)) = k \oplus k \oplus m = m$ and so the one-time pad does constitute a valid encryption scheme.

One can easily prove perfect secrecy of the one-time pad using Lemma 2.4 and the fact that the ciphertext is uniformly distributed regardless of what message is encrypted. We give a proof based directly on the original definition.

THEOREM 2.9 *The one-time pad encryption scheme is perfectly secret.*

CONSTRUCTION 2.8

Fix an integer $\ell > 0$. The message space $\mathcal{M}$, key space $\mathcal{K}$, and ciphertext space $\mathcal{C}$ are all equal to $\{0,1\}^\ell$ (the set of all binary strings of length ℓ).

- Gen: the key-generation algorithm chooses a key from $\mathcal{K} = \{0,1\}^\ell$ according to the uniform distribution (i.e., each of the 2^ℓ strings in the space is chosen as the key with probability exactly $2^{-\ell}$).
- Enc: given a key $k \in \{0,1\}^\ell$ and a message $m \in \{0,1\}^\ell$, the encryption algorithm outputs the ciphertext $c := k \oplus m$.
- Dec: given a key $k \in \{0,1\}^\ell$ and a ciphertext $c \in \{0,1\}^\ell$, the decryption algorithm output the message $m := k \oplus c$.

The one-time pad encryption scheme.

PROOF We first compute $\Pr[\mathsf{Enc}_K(m') = c]$ for arbitrary $m' \in \mathcal{M}$. For the one-time pad,

$$\Pr[\mathsf{Enc}_K(m') = c] = \Pr[m' \oplus K = c] = \Pr[K = m' \oplus c] = 1/2^\ell\,,$$

where the final equality holds because the key is a uniform ℓ-bit string. Fix any distribution over $\mathcal{M}$, any $m \in \mathcal{M}$ with $\Pr[M = m] \neq 0$, and any $c \in \mathcal{C}$. Note that

$$\begin{aligned}
\Pr[C = c] &= \sum_{m' \in \mathcal{M}} \Pr[C = c \mid M = m'] \cdot \Pr[M = m'] \\
&= \sum_{m' \in \mathcal{M}} \Pr[\mathsf{Enc}_K(m') = c] \cdot \Pr[M = m'] \\
&= 2^{-\ell} \cdot \sum_{m' \in \mathcal{M}} Pr[M = m'] \;=\; 2^{-\ell} \;\neq\; 0,
\end{aligned}$$

where the sum is over $m' \in \mathcal{M}$ with $\Pr[M = m'] \neq 0$. Using Bayes' Theorem:

$$\begin{aligned}
\Pr[M = m \mid C = c] &= \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\Pr[C = c]} \\
&= \frac{2^{-\ell} \cdot \Pr[M = m]}{2^{-\ell}} \\
&= \Pr[M = m].
\end{aligned}$$

We conclude that the one-time pad is perfectly secret. ■

The one-time pad was used by several national intelligence agencies in the mid-20th century to encrypt sensitive traffic. Perhaps most famously, the "red phone" linking the White House and the Kremlin during the Cold War was protected using one-time pad encryption, where the governments of the US and USSR would exchange extremely long keys using trusted couriers carrying briefcases of paper on which random characters were written.

Notwithstanding the above, one-time pad encryption is rarely used anymore due to a number of drawbacks it has. Most prominent is that *the key is as long as the message*.[2] This limits applicability of the scheme for sending very long messages (as it may be difficult to securely store a very long key), and is problematic when the parties cannot predict in advance (an upper bound on) how long the message will be.

Moreover, the one-time pad—as the name indicates—*is only "secure" if used once* (with the same key). (It follows that it cannot possibly be secure against stronger attacks such as chosen-plaintext attacks.) Although we did not yet define a notion of security when multiple messages are encrypted, it is easy to see that encrypting more than one message leaks a lot of information. In particular, say two messages m, m' are encrypted using the same key k. An adversary who obtains $c = m \oplus k$ and $c' = m' \oplus k$ can compute

$$c \oplus c' = (m \oplus k) \oplus (m' \oplus k) = m \oplus m'$$

and thus learn the exclusive-or of the two messages or, equivalently, exactly where the two messages differ. While this may not seem very significant, it is enough to rule out any claims of perfect secrecy for encrypting two messages. Moreover, if the messages correspond to natural-language text, then given the exclusive-or of two sufficiently long messages it is possible to perform frequency analysis (as in the previous chapter, though more complex) and recover the messages themselves. An interesting historical example of this is given by the *VENONA project*, as part of which the US and UK were able to decrypt ciphertexts sent by the Soviet Union that were mistakenly encrypted with repeated portions of a one-time pad over several decades.

2.3 Limitations of Perfect Secrecy

We ended the previous section by noting some drawbacks of the one-time pad encryption scheme. Here, we show that these drawbacks are not specific to that scheme, but are instead *inherent* limitations of perfect secrecy. Specifically, we prove that *any* perfectly secret encryption scheme must have a key space that is at least as large as the message space. If all keys are the same length, and the message space consists of all strings of some fixed length, this implies that the key is at least as long as the message. In particular, the key length of the one-time pad is optimal. (The other limitation—namely, that the key can be used only once—is also inherent if perfect secrecy is required; see Exercise 2.10.)

[2]This does not make the one-time pad useless, since it may be easier for two parties to share a key at some point in time before the message to be communicated is known.

THEOREM 2.10 *If* (Gen, Enc, Dec) *is a perfectly secret encryption scheme with message space* $\mathcal{M}$ *and key space* $\mathcal{K}$, *then* $|\mathcal{K}| \geq |\mathcal{M}|$.

PROOF We show that if $|\mathcal{K}| < |\mathcal{M}|$ then the scheme cannot be perfectly secret. Assume $|\mathcal{K}| < |\mathcal{M}|$. Consider the uniform distribution over $\mathcal{M}$ and let $c \in \mathcal{C}$ be a ciphertext that occurs with non-zero probability. Let $\mathcal{M}(c)$ be the set of all possible messages that are possible decryptions of c; that is

$$\mathcal{M}(c) \stackrel{\text{def}}{=} \{m \mid m = \mathsf{Dec}_k(c) \text{ for some } k \in \mathcal{K}\}.$$

Clearly $|\mathcal{M}(c)| \leq |\mathcal{K}|$. (Recall that we may assume Dec is deterministic.) If $|\mathcal{K}| < |\mathcal{M}|$, there is some $m' \in \mathcal{M}$ such that $m' \notin \mathcal{M}(c)$. But then

$$\Pr[M = m' \mid C = c] = 0 \neq \Pr[M = m'],$$

and so the scheme is not perfectly secret. ■

Perfect secrecy with shorter keys? The above theorem shows an inherent limitation of schemes that achieve perfect secrecy. Even so, individuals occasionally claim they have developed a radically new encryption scheme that is "unbreakable" and achieves the security of the one-time pad without using keys as long as what is being encrypted. The above proof demonstrates that such claims *cannot* be true; anyone making such claims either knows very little about cryptography or is blatantly lying.

2.4 * Shannon's Theorem

In his work on perfect secrecy, Shannon also provided a characterization of perfectly secret encryption schemes. This characterization says that, under certain conditions, the key-generation algorithm Gen must choose the key *uniformly* from the set of all possible keys (as in the one-time pad); moreover, for every message m and ciphertext c there is a *unique* key mapping m to c (again, as in the one-time pad). Beyond being interesting in its own right, this theorem is a useful tool for proving (or disproving) perfect secrecy of suggested schemes. We discuss this further after the proof.

The theorem as stated here assumes $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$, meaning that the sets of plaintexts, keys, and ciphertexts are all of the same size. We have already seen that, for perfect secrecy, we must have $|\mathcal{K}| \geq |\mathcal{M}|$. It is easy to see that correct decryption requires $|\mathcal{C}| \geq |\mathcal{M}|$. Therefore, in some sense, perfectly secret encryption schemes with $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$ are "optimally efficient."

THEOREM 2.11 (Shannon's theorem) *Let* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *be an encryption scheme with message space* $\mathcal{M}$, *for which* $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$. *The scheme is perfectly secret if and only if:*

1. *Every key* $k \in \mathcal{K}$ *is chosen with (equal) probability* $1/|\mathcal{K}|$ *by algorithm* Gen.
2. *For every* $m \in \mathcal{M}$ *and every* $c \in \mathcal{C}$, *there exists a unique key* $k \in \mathcal{K}$ *such that* $\mathsf{Enc}_k(m)$ *outputs* c.

PROOF Observe that Enc may as well be deterministic. Otherwise, there is some key k and some message m for which $\mathsf{Enc}_k(m)$ can output at least two possible ciphertexts c_1, c_2. Since $|\mathcal{C}| = |\mathcal{M}|$, there must be some other message $m' \neq m$ for which $\mathsf{Enc}_k(m')$ can output a ciphertext in $\{c_1, c_2\}$. But then correctness fails to hold.

We first prove that if the encryption scheme satisfies conditions 1 and 2, then it is perfectly secret. The proof is essentially the same as the proof of perfect secrecy for the one-time pad, so we will be relatively brief. Fix arbitrary $c \in \mathcal{C}$ and $m \in \mathcal{M}$. Let k be the unique key, guaranteed by condition 2, for which $\mathsf{Enc}_k(m) = c$. Then

$$\Pr[\mathsf{Enc}_K(m) = c] = \Pr[K = k] = 1/|\mathcal{K}|,$$

where the final equality holds by condition 1. So

$$\Pr[C = c] = \sum_{m \in \mathcal{M}} \Pr[\mathsf{Enc}_K(m) = c] \cdot \Pr[M = m] = 1/|\mathcal{K}|.$$

Fixing any distribution over $\mathcal{M}$, any $m \in \mathcal{M}$ with $\Pr[M = m] \neq 0$, and any $c \in \mathcal{C}$, we have:

$$\begin{aligned}
\Pr[M = m \mid C = c] &= \frac{\Pr[C = c \mid M = m] \cdot \Pr[M = m]}{\Pr[C = c]} \\
&= \frac{\Pr[\mathsf{Enc}_K(m) = c] \cdot \Pr[M = m]}{\Pr[C = c]} \\
&= \frac{|\mathcal{K}|^{-1} \cdot \Pr[M = m]}{|\mathcal{K}|^{-1}} \;=\; \Pr[M = m],
\end{aligned}$$

and the scheme is perfectly secret.

For the second direction, assume the encryption scheme is perfectly secret; we show that conditions 1 and 2 hold. Fix arbitrary $c \in \mathcal{C}$. There must be some message m for which $\Pr[\mathsf{Enc}_K(m) = c] \neq 0$. Lemma 2.4 then implies that $\Pr[\mathsf{Enc}_K(m) = c] \neq 0$ for every $m \in \mathcal{M}$. In other words, if we let $\mathcal{M} = \{m_1, m_2, \ldots\}$, then for each $m_i \in \mathcal{M}$ we have a non-empty set of keys $\mathcal{K}_i \subset \mathcal{K}$ such that $\mathsf{Enc}_k(m_i) = c$ if and only if $k \in \mathcal{K}_i$. Moreover, when $i \neq j$ then $\mathcal{K}_i$ and $\mathcal{K}_j$ must be disjoint or else correctness fails to hold. Since

$|\mathcal{K}| = |\mathcal{M}|$, we see that each $\mathcal{K}_i$ contains only a single key k_i, as required by condition 2. Now, Lemma 2.4 shows that for any $m_i, m_j \in \mathcal{M}$ we have

$$\Pr[K = k_i] = \Pr[\mathsf{Enc}_K(m_i) = c] = \Pr[\mathsf{Enc}_K(m_j) = c] = \Pr[K = k_j].$$

Since this holds for all $1 \leq i, j \leq |\mathcal{M}| = |\mathcal{K}|$, and $k_i \neq k_j$ for $i \neq j$, this means each key is chosen with probability $1/|\mathcal{K}|$ as required by condition 1. ■

Uses of Shannon's theorem. Shannon's theorem is useful for deciding whether a given scheme is perfectly secret. Condition 1 is easy to check, and condition 2 can be demonstrated (or contradicted) without having to compute any probabilities (in contrast to working with Definition 2.3 directly). As an example, perfect secrecy of the one-time pad is trivial to prove using Shannon's theorem. We stress, however, that Theorem 2.11 can only be applied when $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$.

References and Additional Reading

The one-time pad is popularly credited to Vernam [170] (who filed a patent for it), but recent historical research [32] shows that it was invented some 35 years earlier. Analysis of the one-time pad had to await the ground-breaking work of Shannon [155], who introduced the notion of perfect secrecy.

In this chapter we studied perfectly secure *encryption*. Some other cryptographic problems can also be solved with "perfect" security. A notable example is the problem of message authentication where the aim is to prevent an adversary from (undetectably) modifying a message sent from one party to another. We study this problem in depth in Chapter 4, discussing "perfectly secure" message authentication in Section 4.6.

Exercises

2.1 Prove that, by redefining the key space, we may assume the key-generation algorithm Gen chooses a key uniformly at random, without changing $\Pr[C = c \mid M = m]$ for any m, c.

> **Hint:** Define the key space to be the set of all possible random tapes for the randomized algorithm Gen.

2.2 Prove that, by redefining the key space, we may assume that Enc is deterministic without changing $\Pr[C = c \mid M = m]$ for any m, c.

2.3 Prove or refute: An encryption scheme with message space $\mathcal{M}$ is perfectly secret if and only if for every probability distribution over $\mathcal{M}$ and every $c_0, c_1 \in \mathcal{C}$ we have $\Pr[C = c_0] = \Pr[C = c_1]$.

2.4 Prove the second direction of Lemma 2.4.

2.5 Prove Lemma 2.6.

2.6 When using the one-time pad with the key $k = 0^\ell$, we have $\mathsf{Enc}_k(m) = k \oplus m = m$ and the message is sent in the clear! It has therefore been suggested to modify the one-time pad by only encrypting with $k \neq 0^\ell$ (i.e., to have Gen choose k uniformly at random from the set of *non-zero* keys of length ℓ). Is this modified scheme still perfectly secret? Explain.

2.7 Let Π denote the Vigenère cipher where the message space consists of all 3-character strings, and the key is generated by first choosing the period t uniformly from $\{1, 2, 3\}$ and then letting the key be a uniform string of length t.

(a) Define $\mathcal{A}$ as follows: $\mathcal{A}$ outputs $m_0 = \texttt{aab}$ and $m_1 = \texttt{abb}$. When given a ciphertext c, it outputs 0 if the first character of c is the same as the second character of c, and outputs 1 otherwise. Compute $\Pr[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1]$.

(b) Construct and analyze an adversary $\mathcal{A}'$ for which $\Pr[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A}',\Pi} = 1]$ is greater than your answer from part (a).

2.8 In this exercise, we look at different conditions under which the shift, mono-alphabetic substitution, and Vigenère ciphers are perfectly secret:

(a) Prove that if only a single character is encrypted, then the shift cipher is perfectly secret.

(b) What is the largest message space $\mathcal{M}$ for which the mono-alphabetic substitution cipher provides perfect secrecy?

(c) Prove that the Vigenère cipher using (fixed) period t is perfectly secret when used to encrypt messages of length t.

Reconcile this with the attacks shown in the previous chapter.

2.9 Prove that a scheme satisfying Definition 2.5 must have $|\mathcal{K}| \geq |\mathcal{M}|$ without using Lemma ??. Specifically, let Π be an arbitrary encryption scheme with $|\mathcal{K}| < |\mathcal{M}|$. Show an $\mathcal{A}$ for which $\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] > \frac{1}{2}$.

Hint: It may be easier to let $\mathcal{A}$ be randomized.

2.10 Consider the following definition of perfect secrecy for the encryption of two messages. Encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ with message space $\mathcal{M}$ is *perfectly secret for two messages* if for all distributions over $\mathcal{M} \times \mathcal{M}$

(i.e., distributions over *pairs* of messages), all $m, m' \in \mathcal{M}$, and all $c, c' \in \mathcal{C}$ with $\Pr[C = c \wedge C' = c'] > 0$:

$$\Pr\left[M = m \wedge M' = m' \mid C = c \wedge C' = c'\right] = \Pr[M = m \wedge M' = m'],$$

where a *single* key, output by Gen, is used to encrypt both m and m'. Prove that *no* encryption scheme satisfies this definition.

Hint: Take $m \neq m'$ but $c = c'$.

2.11 Consider the following definition of perfect secrecy for the encryption of two messages. Encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ with message space $\mathcal{M}$ is *perfectly secret for two different messages* if for all distributions over $\mathcal{M} \times \mathcal{M}$ where the first and second messages are guaranteed to be different (i.e., distributions over *pairs of distinct* messages), all $m, m' \in \mathcal{M}$, and all $c, c' \in \mathcal{C}$ with $\Pr[C = c \wedge C' = c'] > 0$:

$$\Pr[M = m \wedge M' = m' \mid C = c \wedge C' = c'] = \Pr[M = m \wedge M' = m']$$

where a *single* key, output by Gen, is used to encrypt both m and m'. Show an encryption scheme that provably satisfies this definition.

Hint: The encryption scheme you propose need not be efficient, though an efficient solution is possible.

2.12 Assume we require only that an encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ with message space $\mathcal{M}$ satisfy the following: For all $m \in \mathcal{M}$, we have $\Pr[\mathsf{Dec}_K(\mathsf{Enc}_K(m)) = m] \geq 2^{-t}$. (This probability is taken over choice of k as well as any randomness used during encryption.) Show that perfect secrecy can be achieved with $|\mathcal{K}| < |\mathcal{M}|$ when $t \geq 1$. Prove a lower bound on the size of $\mathcal{K}$ in terms of t.

2.13 Let $\varepsilon \geq 0$ be a constant. Say an encryption scheme is *ε-perfectly secret* if for every adversary $\mathcal{A}$ it holds that

$$\Pr\left[\mathsf{PrivK}^{\mathsf{eav}}_{\mathcal{A},\Pi} = 1\right] \leq \frac{1}{2} + \varepsilon\,.$$

(Compare to Definition 2.5.) Show that ε-perfect secrecy can be achieved with $|\mathcal{K}| < |\mathcal{M}|$ when $\varepsilon > 0$. Prove a lower bound on the size of $\mathcal{K}$ in terms of ε.