Textbook: Introduction to Cryptography 2nd ed.
By J.A. Buchmann

Chap 4 Probability and Perfect Secrecy

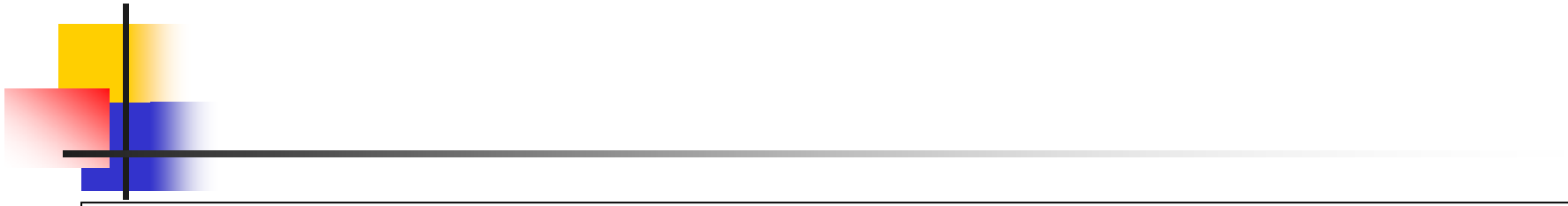
Department of Computer Science and Information Engineering,
Chaoyang University of Technology
朝陽科技大學資工系

Speaker: Fuw-Yi Yang 楊伏夷

伏夷非征番,
道德經 察政章(Chapter 58) **伏**者潛藏也
道紀章(Chapter 14) 道無形象, 視之不可見者曰**夷**

Contents

4.1 Probability



Let S be a finite nonempty set.
We call it the ***sample space***.
Its elements are called ***elementary events***.
The elementary events **model** outcomes of experiments.

Example 4.1.1
If we flip a coin, we either obtain heads H or tails T.
The ***sample space*** is $S = \{H, T\}$.

If we throw a die, then we obtain a number in $\{1, 2, 3, 4, 5, 6\}$.
Therefore, the ***sample space*** is $S = \{1, 2, 3, 4, 5, 6\}$.

4.1 Probability

An ***event*** (for S) is a **subset** of the **sample space *S***.
The certain event is the set ***S*** itself.
The null event is the empty set $\varnothing$.
We say that two events A and B are *mutually exclusive* if their intersection is empty.
The set of all events is the power set $P(S)$ of S.

Example 4.1.2 An **event** is, for example, to obtain an even number when throwing a die. Formally, this event is $\{2, 4, 6\}$. It excludes the event $\{1, 3, 5\}$ to obtain an odd number.

Note: **event** is a set (subset of sample space);
elementary event is an element of sample space.

4.1 Probability

A ***probability distribution*** on S is a map **Pr** that sends an event to a real number, namely

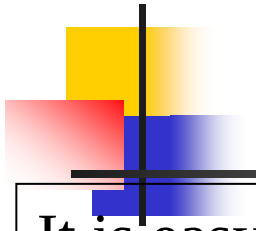
$\Pr: P(S) \to R$,

and has the following properties:

1. $\Pr(A) \geq 0$ for all A,
2. $\Pr(S) = 1$,
3. $\Pr(A \cup B) = \Pr(A) + \Pr(B)$ for two events A and B, which are mutually exclusive.

If A is an **event**, then $\Pr(A)$ is the ***probability*** of this event.
The ***probability*** of an **elementary event** $a \in S$ is $\Pr(a) = \Pr(\{a\})$.

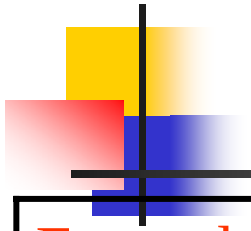
4.1 Probability



It is easy to see that $\Pr(\varnothing) = 0$.
Moreover, $A \subset B$ implies $\Pr(A) \leq \Pr(B)$.
Therefore, $0 \leq \Pr(A) \leq 1$ for all $A \in P(S)$.
Moreover, $\Pr(S \backslash A) = 1 - \Pr(A)$.
If $A_1, \ldots, A_n$ are pairwise mutually exclusive events, then
$\Pr(A_1 \cap \ldots \cap A_n) = \Pr(A_1) + \ldots + \Pr(A_n)$.

Since A is a finite set, it suffices to define the probability distribution on elementary events.
If A is an event then $\Pr(A) = \sum_{a \in A} \Pr(a)$.

4.1 Probability



Example 4.1.3 The probability distribution on the set {1, 2, 3, 4, 5, 6}, which models throwing a die, maps each elementary event to 1/6. The probability of the event "even result" is

$\Pr(\{2, 4, 6\}) = \Pr(2) + \Pr(4) + \Pr(6) = 1/2.$

The probability distribution that maps each elementary event $a \in S$ to the probability $\Pr(a) = 1/|S|$ is called the ***uniform distribution***.

4.2 Conditional Probability

Let S be a ***sample space***, and let Pr be a ***probability distribution*** on S.

Example 4.2.1 Again, we model throwing a die. The sample space is {1, 2, 3, 4, 5, 6}, and Pr sends any elementary event to 1/6.

Suppose Clause has thrown one of the numbers 4, 5, 6, so we know that the *event* $B = \{4, 5, 6\}$ has happened.

Under this assumption, we want to determine the probability that Clause has thrown an *even number*. Each *elementary event* in B is equally probable.

Therefore, each *elementary event* in B has probability 1/3. Since two numbers in B are even, the probability that Clause has thrown an number is 2/3.

4.2 Conditional Probability

Definition 4.2.2 Let A and B be events and $\Pr(B) > 0$.

The ***conditional probability*** of "***A* given that *B***" occurs is defined to be $\Pr(A|B) = \Pr(A \cap B) / \Pr(B)$.

This definition can be understood as follows. The event B will occur a fraction $\Pr(B)$ of the time. Also both A and B will occur a fraction $\Pr(A \cap B)$ of the time.

The ratio $\Pr(A \cap B) / \Pr(B)$ thus gives the *proportion of the time when B occurs, that A also occurs*.

That is , if we **ignores** all the times that B does not occur, and consider only those times that B does occur, then the ratio $\Pr(A \cap B) / \Pr(B)$ equals the fraction of the time that A will occur. This precisely what is the meant by the conditional probability of A given B.

4.2 Conditional Probability

Two events A and B are called *independent* if

$$\Pr(A \cap B) = \Pr(A)\Pr(B).$$

If the events are not independent, we call them *dependent*.

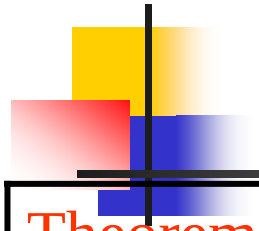
Example 4.2.3 If we flip two coins, then the probability of the event "**the first coin comes up tails**" is **independent** from the event "**the second coin comes up tails**".

The probability that both events occur is1/4. The probability of each individual is1/2.

If the coins are welded together such that they either both fall heads or both tails, then the probability of two tails is $1/2 \neq (1/2) * (1/2)$.

Hence the events "**the first coin comes up tails**" and "**the second coin comes up tails**" are **dependent**.

4.2 Conditional Probability



Theorem 4.2.4 (Theorem of Bayes)

If A and B are events with $\Pr(A) > 0$ and $\Pr(B) > 0$, then

$\Pr(B)\Pr(A|B) = \Pr(A)\Pr(B|A)$.

Proof. $\Pr(A|B) = \Pr(A \cap B) / \Pr(B) \Rightarrow \Pr(A \cap B) = \Pr(B)\Pr(A|B)$

$\Pr(B|A) = \Pr(A \cap B) / \Pr(A) \Rightarrow \Pr(A \cap B) = \Pr(A)\Pr(B|A)$.

end of proof

4.3 Birthday paradox

Birthday paradox: Suppose a group of people are in a room. What is the probability that two of them have the same birthday?

Suppose there are n birthdays and that there are k people in the room. An **elementary event** is a tuple $(b_1, \ldots, b_k) \in \{1, 2, \ldots, n\}^k$.

The birthday of the ith person is b_i, $1 \le i \le k$, so we have n^k elementary events.

We assume that those elementary events are equally probable. Then the probability of an elementary events is $1/ n^k$.

Let p be the probability that two people in the room have the same birthday. Then with probability $q = 1 - p$ any two people have different birthdays. Next page

4.3 Birthday paradox

Let E be the set of all vectors $(g_1, \ldots, g_k) \in \{1, 2, \ldots, n\}^k$ whose entries are *pairwise different*. Then E models the *Birthday paradox*.

Let $|E|$ denote the number of element in E. Then

$$|E| = (n - 0)\,(n - 1)\,\ldots\,(n - (k - 1)) = \prod_{i=0}^{k-1}(n-i) \text{ and}$$

$$q = (n - 0)\,(n - 1)\,\ldots\,(n - (k - 1)) / n^k$$

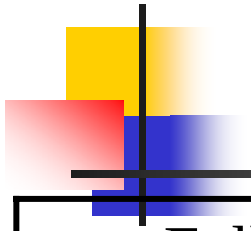
$$= \frac{1}{n^k}\prod_{i=0}^{k-1}(n-i) = \prod_{i=1}^{k-1}(1-\frac{i}{n}) .$$

Since $1 + x \le e^x$ holds for all real numbers, therefore

$$q \le \prod_{i=1}^{k-1} e^{-i/n} = e^{-\sum_{i=1}^{k-1} i/n} = e^{-k(k-1)/2n}.$$

If $k \ge (1 + (1 + 8n \log 2)^{0.5}) / 2$, then $q \le 0.5$.

4.4 Perfect Secrecy



Following Shannon, we will introduce ***perfect secrecy***.

Assume that Alice uses a ***cryptosystem*** to send encrypted messages to Bob. If she sends such an encrypted message to Bob, the attacker, Oscar, can read the ciphertext.

Oscar tries to obtain information concerning the plaintext from the ciphertext.

A cryptosystem has ***perfect secrecy*** if Oscar **learns nothing** about the plaintext from the ciphertext.

We want to formalize this property mathematically.

4.4 Perfect Secrecy

The cryptosystem has a finite **plaintext space $\mathcal{P}$**, a finite **ciphertext space $\mathcal{C}$**, and a finite **key space $\mathcal{K}$**.

The encryption functions are E_k, $k \in \mathcal{K}$ and the decryption functions are D_k, $k \in \mathcal{K}$.

We assume that the probability of a plaintext p is $\Pr_{\mathcal{P}}(p)$. The function $\Pr_{\mathcal{P}}$ is a probability distribution on the plaintext space; the probability of a key k is $\Pr_{\mathcal{K}}(k)$. The function $\Pr_{\mathcal{P}}$ is a probability distribution on the key space.

The probability that a plaintext p occurs and is encrypted with key k is $\Pr(p, k) = \Pr_{\mathcal{P}}(p) \Pr_{\mathcal{K}}(k)$. ---4.4

This defines a probability distribution Pr on the sample space $\mathcal{P} \times \mathcal{K}$

4.4 Perfect Secrecy

Consider the sample space $\mathcal{P}\times\mathcal{K}$.

If p is a plaintext, then we also denote the event $\{(p, k): k \in \mathcal{K}\}$ that p is encrypted. Clearly, we have $\Pr(p) = \Pr_{\mathcal{P}}(p)$.

Also for a key $k \in \mathcal{K}$ we denote by k the event $\{(p, k): p \in \mathcal{K}\}$ that the key k is chosen for encryption. Clearly, we have $\Pr(k) = \Pr_{\mathcal{K}}(k)$.

By eq. 4.4, the events p and k are independent.

For a ciphertext $c \in \mathcal{C}$ we denote by c the event $\{(p, k): E_k(p) = c\}$ that the result of the encryption is c.

4.4 Perfect Secrecy

Definition 4.4.1 The cryptosystem of this section has ***perfect secrecy*** if the events that a *particular ciphertext* occurs and that a *particular plaintext* has been encrypted are independent. Namely, $\Pr(p|c) = \Pr(p)$ for all plaintexts p and all ciphertext c.

Example 4.4.2 Let $\mathcal{P} = \{0, 1\}$, $\Pr(0) = 1/4$, $\Pr(1) = 3/4$.

Also, let $\mathcal{K} = \{A, B\}$, $\Pr(A) = 1/4$, $\Pr(B) = 3/4$.

Finally, Let $\mathcal{C} = \{a, b\}$.

Then the probability that the plaintext 1 occurs and is encrypted with key B is $\Pr(1)\ \Pr(B) = 9/16$.

The encryption function E_k works as follows:

$E_A(0) = a$, $E_A(1) = b$, $E_B(0) = b$, $E_B(1) = a$. *see next page*

4.4 Perfect Secrecy

$\Pr(0) = 1/4$, $\Pr(1) = 3/4$, $\Pr(A) = 1/4$, $\Pr(B) = 3/4$.

$EA(0) = a$, $\mathbf{EA(1) = b}$, $\mathbf{EB(0) = b}$, $EB(1) = a$.

The probability of the ciphertext a is

$$\Pr(a) = \Pr(0, A) + \Pr(1, B) = 1/16 + 9/16 = 10/16.$$

Similarly, $\Pr(b) = \Pr(1, A) + \Pr(0, B) = 3/16 + 3/16 = 6/16$.

Now computes the conditional probability $\Pr(p|c)$ for all plaintexts p and all ciphertexts c.

$\Pr(0|a) = \Pr(\ \{0\}\cap\{a\}\ ) / \Pr(a) = (1/4)(1/4)/(10/16) = 1/10$

$\Pr(1|a) = \Pr(\ \{1\}\cap\{a\}\ ) / \Pr(a) = (3/4)(3/4)/(10/16) = 9/10$

$\Pr(0|b) = \Pr(\ \mathbf{\{0\}\cap\{b\}}\ ) / \Pr(b) = (1/4)(3/4)/(6/16) = 3/6$

$\Pr(1|b) = \Pr(\ \mathbf{\{1\}\cap\{b\}}\ ) / \Pr(b) = (3/4)(1/4)/(6/16) = 3/6$

Those results show that the cryptosystem described does not have perfect secrecy. end of example

4.4 Perfect Secrecy

Theorem 4.4.3 Let $|\mathcal{P}| = |\mathcal{K}| = |C| < \infty$ and $\Pr(p) > 0$ for any plaintext p.

Our cryptosystem has ***perfect secrecy*** if and only if the probability distribution on the key space is the uniform distribution and if for any plaintext p and any ciphertext c there is exactly one key k with $E_k(p) = c$.

Proof……

Example 4.4.4 *see next page*

4.4 Perfect Secrecy

Example 4.4.4 Theorem 4.4.3 implies that the cryptosystem from example 4.4.2 has perfect secrecy if we set $\Pr(A) = \Pr(B) = 1/2$.

Let $\mathcal{P} = \{0, 1\}$, $\Pr(0) = 1/4$, $\Pr(1) = 3/4$,

$\mathcal{K} = \{A, B\}$, $\Pr(A) = 1/2$, $\Pr(B) = 1/2$, $\mathcal{C} = \{a, b\}$.

The encryption function E_k works as follows:

$E_A(0) = a$, $E_A(1) = b$, $E_B(0) = b$, $E_B(1) = a$.

The probability of the ciphertext a is

$$\Pr(a) = \Pr(0, A) + \Pr(1, B) = 1/8 + 3/8 = 4/8.$$

Similarly, $\Pr(b) = \Pr(1, A) + \Pr(0, B) = 3/8 + 1/8 = 4/8$.

4.4 Perfect Secrecy

$\mathcal{P} = \{0, 1\}$, Pr(0) = 1/4, Pr(1) = 3/4,

$\mathcal{K} = \{A, B\}$, Pr(A) = 1/2, Pr(B) = 1/2,

$E_A(0) = a$, $\boldsymbol{E_A(1) = b}$, $\boldsymbol{E_B(0) = b}$, $E_B(1) = a$.

Pr(0|a) = Pr($\{0\} \cap \{a\}$) / Pr(a) = (1/4)(1/2)/(4/8) = 1/4 = Pr(0)

Pr(1|a) = Pr($\{1\} \cap \{a\}$) / Pr(a) = (3/4)(1/2)/(4/8) = 3/4 = Pr(1)

Pr(0|b) = Pr($\boldsymbol{\{0\} \cap \{b\}}$) / Pr(b) = (1/4)(1/2)/(4/8) = 1/4 = Pr(0)

Pr(1|b) = Pr($\boldsymbol{\{1\} \cap \{b\}}$) / Pr(b) = (3/4)(1/2)/(4/8) = 3/4 = Pr(1)

Those results show that the cryptosystem described does have perfect secrecy. end of example

4.5 Vernam One-Time Pad (patented by Gilbert Vernam in 1917)

The most famous cryptosystem that has perfect secrecy is the ***Vernam one-time pad***.

Let n be a positive integer. The Vernam one-time pad encrypts bitstrings of length n. Plaintext space, ciphertext space, and key space are $\mathcal{P} = \mathcal{K} = C = \{0, 1\}^n$. The encryption function for key $k \in \{0, 1\}^n$ is

$$E_k: \{0, 1\}^n \to \{0, 1\}^n, p \to p \oplus k.$$

The decryption function for key k is the same.

To encrypt a plaintext $p \in \{0, 1\}^n$, Alice chooses a key k ***randomly with uniform distribution*** from the set $\{0, 1\}^n$. She compute the ciphertext $c = p \oplus k$. By Theorem 4.4.3, this cryptosystem is **perfectly secure** because the uniform distribution is used on the key space and for each plaintext p and each ciphertext c there is exactly one key k with $c = p \oplus k$, namely $k = p \oplus c$.

4.5 Vernam One-Time Pad

The following example demonstrates that the One-Time-Pad ***is not secure against active attacks***.

Example 4.5.1 Alice encrypts her electronic bank transactions using the one-time pad. If the attacker Oscar knows where in the ciphertext the amount is encrypted, then he can change that part of the ciphertext.

(Applying integrity check is a countermeasures against such active attack)

4.6 Random numbers

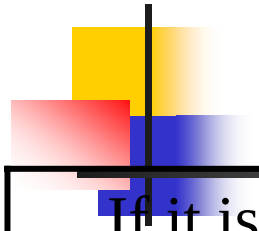
We want to generate ***uniformly distributed random numbers*** in the set $\{0, 1, \ldots, m\}$, $m \in \mathrm{N}$. We set n = size $m = \lfloor \log m \rfloor + 1$.

Then we generate n random bits $b_1, \ldots, b_n$.

If the number $a = b_1 2^{n-1} + \ldots + b_n 2^{n-n}$ is greater than m, then we forget it and generate a new one in the same way. Otherwise, a is a random number.

If we want to generate ***uniformly distributed random n-bit numbers***, $n \in \mathrm{N}$, then we generate n-1 random bits $b_2, \ldots, b_n$ and set $b_1 = 1$ and output $a = b_1 2^{n-1} + \ldots + b_n 2^{n-n}$.

4.6 Pseudorandom numbers



If it is too time-consuming to generate true random numbers, then ***pseudorandom number*** generators are used.

A pseudorandom number generator is an algorithm, given a short sequence of random bits, produces a long sequence of bits that looks random.