

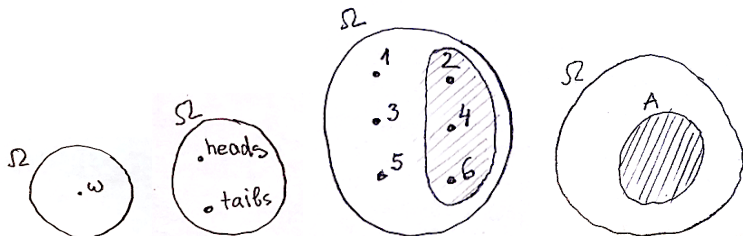
# Theory of Unbreakable Ciphers

Ahto Buldas

September 30, 2019

# Sample Space and Events

$\Omega$ -sample space, that contains all possible outcomes  $\omega \in \Omega$ .



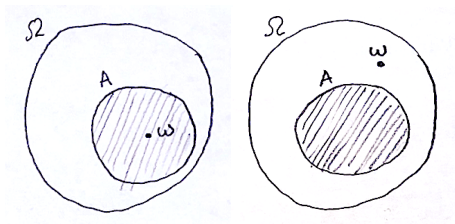
For example,  $\Omega = \{\text{heads}, \text{tails}\}$  for a coin, and  $\Omega = \{1, \dots, 6\}$  for a die.

**Events** are subsets  $A \subseteq \Omega$ .

For a die, the event  $\{2, 4, 6\}$  means that the outcome is even.

# When do Events Happen?

An event  $A$  *happens* if  $\omega \in A$  for the actual outcome  $\omega$ .



Empty event  $\emptyset$  is called the *impossible event* (it *never* happens)

$\Omega$  is called the *universal event* (it *always* happens)

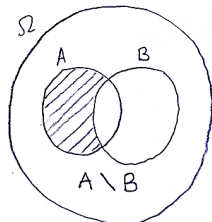
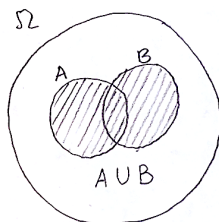
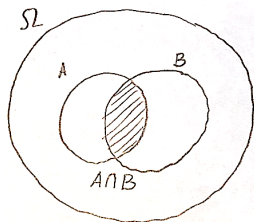
# Operations with Events

For every two events  $A$  and  $B$  we can compute:

*Intersection*     $A$  and  $B$      $A \cap B = \{\omega \in \Omega : \omega \in A \text{ and } \omega \in B\}$

*Union*     $A$  or  $B$      $A \cup B = \{\omega \in \Omega : \omega \in A \text{ or } \omega \in B\}$

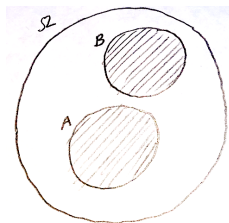
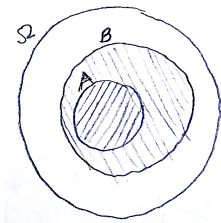
*Difference*     $A$  but not  $B$      $A \setminus B = \{\omega \in \Omega : \omega \in A \text{ and } \omega \notin B\}$



# Relations Between Events

**Inclusion:** Event  $A$  *implies* event  $B$ , if  $A \subseteq B$ , i.e. if  $\omega \in A$  always implies  $\omega \in B$ . If  $A$  happens then  $B$  happens.

**Exclusion:** Events  $A$  and  $B$  are *mutually exclusive* if  $A \cap B = \emptyset$ , i.e.  $A$  and  $B$  cannot simultaneously happen.



# Some Properties

## Theorem (1)

$$A = (A \setminus B) \cup (A \cap B)$$

### Proof.

We prove (a)  $A \subseteq (A \setminus B) \cup (A \cap B)$  and (b)  $(A \setminus B) \cup (A \cap B) \subseteq A$

(a) If  $\omega \in A$  then either:

- $\omega \in B$ , which implies  $\omega \in A \cap B$ , or
- $\omega \notin B$ , which implies  $\omega \in A \setminus B$

(b) If  $\omega \in (A \setminus B) \cup (A \cap B)$ , then either:

- $\omega \in A \setminus B$ , which implies  $\omega \in A$ , or
- $\omega \in A \cap B$ , which also implies  $\omega \in A$



## Some Properties

### Theorem (2)

$$A \cup B = (A \setminus B) \cup B$$

### Proof.

We prove (a)  $A \cup B \subseteq (A \setminus B) \cup B$  and (b)  $(A \setminus B) \cup B \subseteq A \cup B$

(a) If  $\omega \in A \cup B$ , then either:

- $\omega \in B$  or
- $\omega \notin B$  and  $\omega \in A$ , which implies  $\omega \in A \setminus B$ .

(b) If  $\omega \in (A \setminus B) \cup B$  then either:

- $\omega \in B$  or
- $\omega \in A \setminus B$  that implies  $\omega \in A$ .



# Event Algebra

The set  $\mathcal{F}$  of all events we consider must be a *sigma-algebra*:

- $\Omega \in \mathcal{F}$
- If  $A \in \mathcal{F}$ , then  $\Omega \setminus A \in \mathcal{F}$
- If  $A_1, A_2, A_3, \dots \in \mathcal{F}$ , then  $A_1 \cup A_2 \cup A_3 \cup \dots \in \mathcal{F}$

If  $A \in \mathcal{F}$ , then  $A$  is said to be a *measurable* subset.

*Example:* The *set*  $P(\Omega)$  *of all subsets* of  $\Omega$  is a sigma-algebra.

In this class, we mostly assume that  $\mathcal{F} = P(\Omega)$ .



# Probability Measure

*Probability (measure)* is a function  $P: \mathcal{F} \rightarrow \mathbb{R}$  such that:

- *PM1*:  $0 \leq P[A] \leq 1$  for any event  $A \in \mathcal{F}$ .
- *PM2*:  $P[\Omega] = 1$
- *PM3*: If  $A_1, A_2, \dots \in \mathcal{F}$  are mutually exclusive, then

$$P[A_1 \cup A_2 \cup \dots] = P[A_1] + P[A_2] + \dots$$

The triple  $(\Omega, \mathcal{F}, P)$  is called a *probability space*.

If  $\mathcal{F}$  is the set of all subsets of  $\Omega$ , we omit  $\mathcal{F}$  and say that a probability space is a pair  $(\Omega, P)$ .

# Some Implications

## Theorem

$$P[\Omega \setminus A] = 1 - P[A]$$

## Proof.

By *PM2*, we have  $P[\Omega] = 1$ . As  $A$  and  $\Omega \setminus A$  are mutually exclusive, and  $(\Omega \setminus A) \cup A = \Omega$ , by *PM3*, we have  $P[\Omega \setminus A] + P[A] = P[\Omega] = 1$  and hence

$$P[\Omega \setminus A] = \underbrace{P[\Omega \setminus A] + P[A]}_1 - P[A] = 1 - P[A] .$$



## Some Implications

### Theorem

$$P[A] + P[B] = P[A \cap B] + P[A \cup B]$$

### Proof.

By Thm. 1:  $A = (A \setminus B) \cup (A \cap B)$ . As  $A \setminus B$  and  $A \cap B$  are mutually exclusive, by **PM3**:  $P[A] = P[A \setminus B] + P[A \cap B]$ . Hence,

$$P[A] + P[B] = P[A \setminus B] + P[B] + P[A \cap B]$$

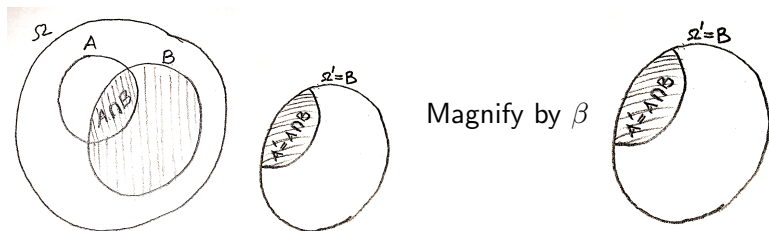
By Thm. 2:  $A \cup B = (A \setminus B) \cup B$ . As  $A \setminus B$  and  $B$  are mutually exclusive, by **PM3**:  $P[A \cup B] = P[A \setminus B] + P[B]$ . Hence,

$$P[A] + P[B] = \underbrace{P[A \setminus B] + P[B]}_{P[A \cup B]} + P[A \cap B] = P[A \cup B] + P[A \cap B] .$$



# Learning

Somehow we learn that an event  $B$  (with  $P[B] \neq 0$ ) happens, i.e.  $\omega \in B$ . Probability space  $(\Omega, P)$  collapses to a new space  $(\Omega', P')$ , where  $\Omega' = B$ .



We want that there is  $\beta$ , so that  $P'[A] = \beta \cdot P[A \cap B]$  for any event  $A$ .  
 As in the new space,  $P'[B] = P'[\Omega'] = 1$ , we have  $\beta = \frac{1}{P[B \cap B]} = \frac{1}{P[B]}$ , i.e.

$$P'[A] = \frac{P[A \cap B]}{P[B]} .$$

# Conditional Probability

The probability

$$P'[A] = \frac{P[A \cap B]}{P[B]}$$

is denoted by  $P[A | B]$  and is called the *conditional probability* of  $A$  assuming that  $B$  happens, i.e.

$$P[A | B] = \frac{P[A \cap B]}{P[B]}$$

*Corollary (Chain Rule):*

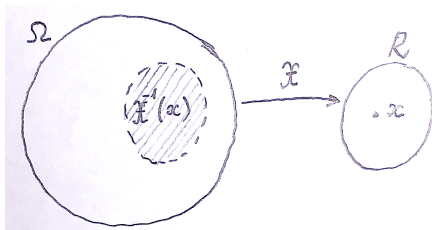
$$P[A \cap B] = P[B] \cdot P[A | B] = P[A] \cdot P[B | A]$$

# Random Variables

*Random variable*  $X$  is any function  $X: \Omega \rightarrow R$ , where  $R$  is called the *range* of  $X$ . We write  $R_X$  to denote the range of  $X$

For any  $x \in R$ , we define  $X^{-1}(x)$  as the event  $\{\omega: X(\omega) = x\}$  and use the notation:

$$P_X[x] = P[X = x] = P[X^{-1}(x)] .$$



# Finite Range Random Variables

In cryptography, we mostly assume that the range  $R$  is *finite*.

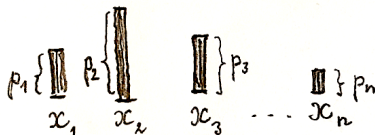
Note that if  $x \neq x'$ , then the events  $X^{-1}(x)$  and  $X^{-1}(x')$  are mutually exclusive and as  $\cup_{x \in R} X^{-1}(x) = \Omega$ , we have:

$$\sum_x \mathbb{P}_X[x] = \mathbb{P}[\cup_{x \in R} X^{-1}(x)] = \mathbb{P}[\Omega] = 1 .$$

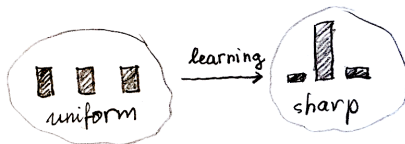
# Probability Distributions and Histograms

Assume  $R$  is finite and  $R = \{x_1, x_2, \dots, x_n\}$ .

The sequence of values  $(p_1, p_2, \dots, p_n)$ , where  $p_i = P_X[x_i]$ , is called the *probability distribution* of  $X$ .



*Histograms* are graphical representations of probability distributions.





## Independent Events and Random Variables

Events  $A$  and  $B$  are said to be *independent* if  $P[A \cap B] = P[A] \cdot P[B]$

If  $P[A] \neq 0 \neq P[B]$ , then independence is equivalent to:

$$P[A | B] = P[A] \quad \text{and} \quad P[B | A] = P[B] ,$$

i.e. the probability of  $A$  does not change, if we learn that  $B$  happened.

We say that  $X$  and  $Y$  are *independent random variables* if for every  $x \in R_X$  and  $y \in R_Y$  :

$$\begin{aligned} P[X = x, Y = y] &= P[X^{-1}(x) \cap Y^{-1}(y)] = P[X^{-1}(x)] \cdot P[Y^{-1}(y)] \\ &= P[X = x] \cdot P[Y = y] . \end{aligned}$$

This means that the probability distribution of  $X$  does not change, if we learn the value of  $Y$ , and vice versa

# Direct Product of Random Variables

By the *direct product*  $XY$  (or  $(X, Y)$ ) of random variables  $X$  and  $Y$  on a probability space  $(\Omega, \mathcal{P})$  is a random variable defined by

$$(XY)(\omega) = (X(\omega), Y(\omega)) .$$

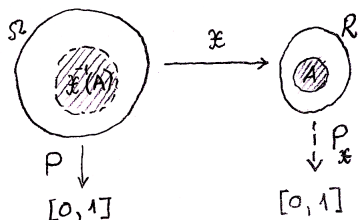
# Factor Space

Let  $X$  be a random variable (with range  $R$ ) on a probability space  $(\Omega, \mathcal{P})$ .

If we take  $\Omega' = R$  and define a probability function  $\mathcal{P}_X$  on  $R$  as follows:

$$\mathcal{P}_X[A] = \mathcal{P}[X^{-1}(A)]$$

where  $X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\}$ , we get a probability space  $(R, \mathcal{P}_X)$  that we call a *factor space*.



# Probabilistic Model of a Cipher

Plaintext  $X$ , key  $Z$  and ciphertext  $Y = E_Z(X)$  are random variables on  $(\Omega, \mathcal{P})$ . It is mostly assumed that  $X$  and  $Z$  are independent.

As we need only  $X$ ,  $Y$ , and  $Z$ , we study the *factor space*  $(R_{XZ}, P_{XZ})$  that consists of all possible plaintext-key pairs  $(x, z)$ , whereas

$$\mathbf{P}_{XZ}[x, z] = \mathbf{P}[X = x] \cdot \mathbf{P}[Z = z] = p(x) \cdot p(z)$$

$X(x, z) = x$ ,  $Z(x, z) = z$ , and  $Y(x, z) = E_z(x)$ .

## Some Observations

$$\begin{aligned}
 p(y) &= \mathbf{P}_{XZ}[Y = y] = \sum_{x,z} \mathbf{P}[x, z][E_z(x) = y] \\
 &= \sum_x p(x) \sum_z p(z)[E_z(x) = y] \\
 p(x, y) &= \mathbf{P}_{XZ}[X = x, Y = y] = \sum_z \mathbf{P}[x, z][E_z(x) = y] \\
 &= p(x) \sum_z p(z)[E_z(x) = y]
 \end{aligned}$$

Here,  $[A(x, yz)]$  is the so-called *Iverson symbol*:

$$[A(x, y, z)] = \begin{cases} 1 & \text{if } A(x, y, z) \text{ holds} \\ 0 & \text{otherwise} \end{cases}$$

## Definition of Unbreakable Cipher

A cipher is *unbreakable* if ciphertext  $Y$  and plaintext  $X$  are independent.

### Theorem

*If  $Z$  is independent of  $X$ ,  $Z$  is uniformly distributed and for every plaintext  $x$  and for every ciphertext  $y$  there is a unique key  $z$  such that  $E_z(x) = y$ , then the cipher is unbreakable.*

### Proof.

Due to the unique  $z$ , we have  $\sum_z p(z)[E_z(x) = y] = p(z)$ , and thus

$$\begin{aligned} p(x | y) &= \frac{p(x, y)}{p(y)} = \frac{p(x) \sum_z p(z)[E_z(x) = y]}{\sum_x p(x) \sum_z p(z)[E_z(x) = y]} = \frac{p(x)p(z)}{p(z) \sum_x p(x)} \\ &= \frac{p(x)p(z)}{p(z) \cdot 1} = p(x) \end{aligned}$$



# Shift Cipher in Unbreakable

Shift cipher:  $y = E_z(x) = x + z \pmod m$

For every  $x$  and  $y$ , there is one and only one  $z$ , such that  $E_z(x) = y$ :

$$z = y - x \pmod m .$$

Therefore, by the theorem above, shift cipher is unbreakable.

# Redundancy of English

In case of 26-letter alphabet, a single letter contains  $\log_2 26 \approx 4.7$  bits of information.

Random  $n$ -letter sequence contains  $4.7n$  bits of information.

Meaningful english texts contain just about 1.5 bits of information per letter.

There are  $2^{4.7n}$  arbitrary  $n$ -letter sequences,  $2^{1.5n}$  of them meaningful

The probability that a randomly chosen  $n$ -letter sequence is meaningful is:

$$\mu = \frac{2^{1.5n}}{2^{4.7n}} = 2^{-3.2n} .$$



# Exhaustive Key Search

Given a ciphertext  $y$

For all keys  $z$ , check if  $D_z(y)$  is a meaningful text

Success, if there is just one  $z$  for which  $D_z(y)$  is meaningful

# Ideal Cipher Model

For every key  $z$ , the function  $E_z: \mathbf{X} \rightarrow \mathbf{Y}$  is a randomly chosen one-to-one function

This implies that the decryption function  $D_z: \mathbf{Y} \rightarrow \mathbf{X}$  is also a randomly chosen one-to-one function

If  $z_1 \neq z_2$ , then  $X_1 = D_{z_1}(y)$  and  $X_2 = D_{z_2}(y)$  are independent uniformly distributed random variables

# Unicity Distance

*Unicity distance*: message length  $n_0$  for which the plaintext can be derived from the ciphertext via exhaustive key search

Let  $y$  be a ciphertext

Assume there are  $2^k$  possible keys  $z$ , one of which is the right key

The probability that  $D_z(y)$  is meaningful for a fixed wrong key  $z$  is  $\mu = 2^{-3.2n}$

The probability that  $D_z(y)$  is meaningful for any of the wrong keys is bounded by  $(2^k - 1)\mu$  and also by  $2^k\mu = 2^{k-3.2n}$

If  $n > n_0 = \frac{k}{3.2}$ , the success probability of exhaustive search increases rapidly

# Unicity Distance for Substitution Ciphers

The number of keys is  $26!$

Hence,  $k = \log_2(26!) \approx 88.4$

Therefore, the unicity distance is  $n_0 = 88.4/3.2 \approx 28$