Theory of Unbreakable Ciphers

Ahto Buldas Aleksandr Lenin

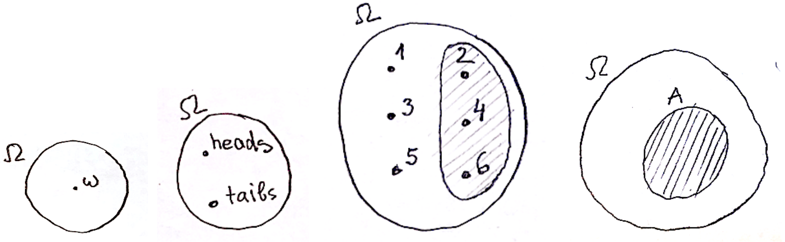
October 1, 2020

Agenda

1. Elementary Probability Theory
2. Unbreakable (Perfect) Ciphers
3. Breaking Imperfect Ciphers

Sample Space and Events

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



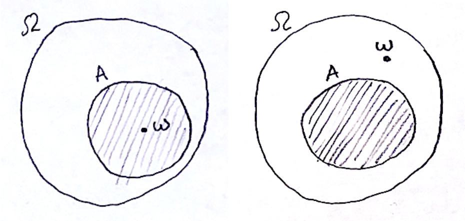
For example, $\Omega = \{\text{heads}, \text{tails}\}$ for a coin, and $\Omega = \{1, \ldots, 6\}$ for a die.

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 ω.



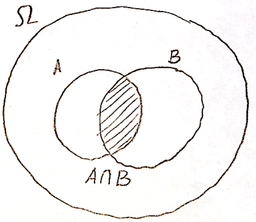
Empty event $\emptyset$ is called the *impossible event* (it *never* happens)

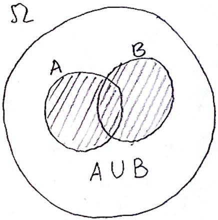
Ω is called the *universal event* (it *always* happens)

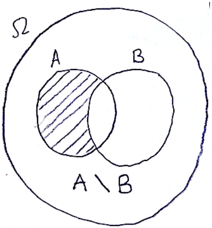
Operations with Events

For every two events A and B we can compute:

Intersection	*A and B*	$A \cap B = \{\omega \in \Omega \colon \omega \in A \text{ and } \omega \in B\}$
Union	*A or B*	$A \cup B = \{\omega \in \Omega \colon \omega \in A \text{ or } \omega \in B\}$
Difference	*A but not B*	$A \setminus B = \{\omega \in \Omega \colon \omega \in A \text{ and } \omega \notin B\}$



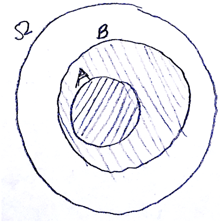


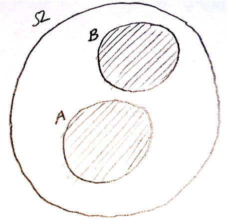


Relations Between Events

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)

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

Proof.

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

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

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

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

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

Some Properties

Theorem (2)

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

Proof.

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

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

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

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

- $\omega \in B$ or
- $\omega \in A \backslash B$ 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 $\Omega \backslash A \in F$
- If $A_1, A_2, A_3, \ldots \in \mathcal{F}$, then $A_1 \cup A_2 \cup A_3 \cup \ldots \in \mathcal{F}$

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

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

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

Probability Measure

Probability (measure) is a function $\mathsf{P}\colon \mathcal{F} \to \mathbb{R}$ such that:

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

$$\mathsf{P}[A_1 \cup A_2 \cup \ldots] = \mathsf{P}[A_1] + \mathsf{P}[A_2] + \ldots$$

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

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

Some Implications

Theorem

$\mathsf{P}[\Omega \backslash A] = 1 - \mathsf{P}[A]$

Proof.

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

$$\mathsf{P}[\Omega \backslash A] = \underbrace{\mathsf{P}[\Omega \backslash A] + \mathsf{P}[A]}_{1} - \mathsf{P}[A] = 1 - \mathsf{P}[A] \ .$$

□

Some Implications

Theorem

$\mathsf{P}[A] + \mathsf{P}[B] = \mathsf{P}[A \cap B] + \mathsf{P}[A \cup B]$

Proof.

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

$$\mathsf{P}[A] + \mathsf{P}[B] = \mathsf{P}[A \backslash B] + \mathsf{P}[B] + \mathsf{P}[A \cap B]$$

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

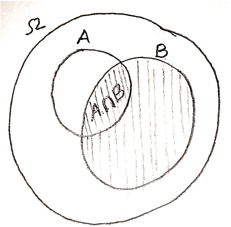
$$\mathsf{P}[A] + \mathsf{P}[B] = \underbrace{\mathsf{P}[A \backslash B] + \mathsf{P}[B]}_{\mathsf{P}[A \cup B]} + \mathsf{P}[A \cap B] = \mathsf{P}[A \cup B] + \mathsf{P}[A \cap B] \ .$$

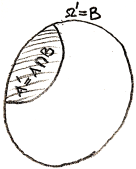
□

Learning

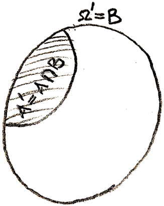
Somehow we learn that an event B (with $\mathsf{P}[B] \neq 0$) happens, i.e. $\omega \in B$.

Probability space (Ω, P) collapses to a new space (Ω', P'), where $\Omega' = B$.





Magnify by β



We want that there is β, so that $\mathsf{P}'[A] = \beta \cdot \mathsf{P}[A \cap B]$ for any event A.

As in the new space, $\mathsf{P}'[B] = \mathsf{P}'[\Omega'] = 1$, we have $\beta = \frac{1}{\mathsf{P}[B \cap B]} = \frac{1}{\mathsf{P}[B]}$, i.e.

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

Conditional Probability

The probability

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

is denoted by $\mathsf{P}[A \mid B]$ and is called the *conditional probability* of A assuming that B happens, i.e.

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

Corollary (Chain Rule):

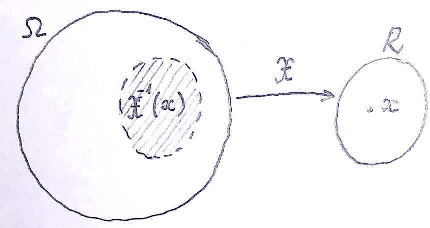
$$\mathsf{P}[A \cap B] = \mathsf{P}[B] \cdot \mathsf{P}[A \,|B] = \mathsf{P}[A] \cdot \mathsf{P}[B \,|A]$$

Random Variables

Random variable X is any function $X\colon \Omega \to R$, 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 $\{\omega\colon X(\omega) = x\}$ and use the notation:

$$\underset{X}{\mathsf{P}}[x] = \mathsf{P}[X = x] = \mathsf{P}[X^{-1}(x)] \ .$$



Finite Range Random Variables

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

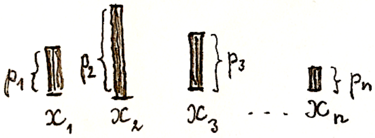
Note that if $x \neq x'$, then the events $X^{-1}(x)$ and $X^{-1}(x')$ are mutually exclusive and as $\cup_{x \in R} X^{-1}(x) = \Omega$, we have:

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

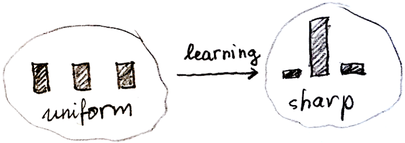
Probability Distributions and Histograms

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

The sequence of values $(p_1, p_2, \ldots, p_n)$, where $p_i = \underset{X}{\mathsf{P}}[x_i]$, 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 $\mathsf{P}[A \cap B] = \mathsf{P}[A] \cdot \mathsf{P}[B]$

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

$$\mathsf{P}[A \mid B] = \mathsf{P}[A] \qquad \text{and} \qquad \mathsf{P}[B \mid A] = \mathsf{P}[B] \ ,$$

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 $y \in R_Y$:

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

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 (Ω, P) is a random variable defined by

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

Factor Space

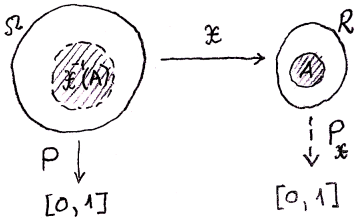
Let $X\colon \Omega \to R$ be a random variable on a probability space (Ω, P).

Then $(R, \underset{X}{\mathsf{P}})$ is also a probability space, where $\underset{X}{\mathsf{P}} = \mathsf{P} \circ X^{-1}$, i.e. $\forall A \subseteq R$:

$$\underset{X}{\mathsf{P}}[A] = \mathsf{P}[X^{-1}(A)]$$

and $X^{-1}(A) = \{\omega \in \Omega\colon X(\omega) \in A\}$.

The space $(R, \underset{X}{\mathsf{P}})$ is called a *factor space*.



Probabilistic Model of a Cipher

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

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

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

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

Some Observations

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

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

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

□

Shift Cipher in Unbreakable

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

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

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

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:

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

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 $E_z\colon \mathbf{X} \to \mathbf{Y}$ is a randomly chosen one-to-one function

This implies that the decryption function $D_z\colon \mathbf{Y} \to \mathbf{X}$ 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 $n_0 = 88.4/3.2 \approx 28$