Information Theory and Coding

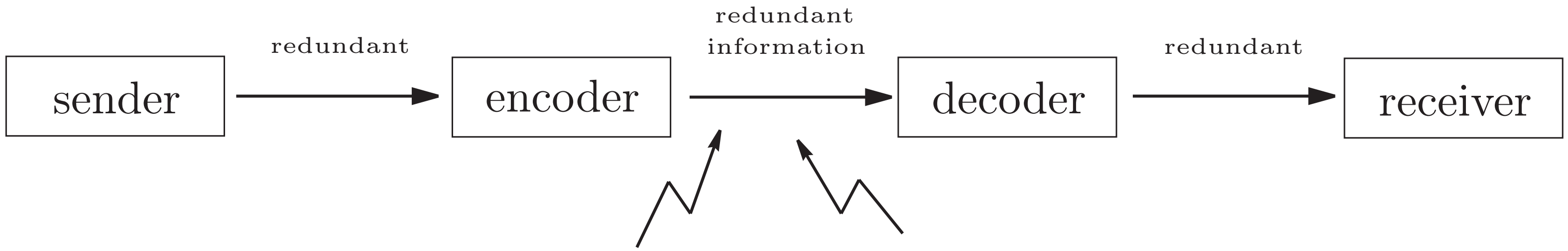
Discrete channels

Cédric RICHARD
Université Côte d'Azur

Channel coding

Motivations

In a real system, the message received by the recipient can differ from that emitted by the source due to perturbations. We talk about *noisy channel*.



Channel coding consists of introducing redundancy into the message

$\longrightarrow$ prevent the loss of information due to the channel

Discrete channel models

General model

A discrete channel is a stochastic system that accepts, as an input, symbol sequences defined on an alphabet $\mathcal{X}$, and outputting sequences of symbols defined on an alphabet $\mathcal{Y}$.

Inputs and outputs are linked by a probabilistic model:

$$P(Y_1 = y_1, \ldots, Y_m = y_m | X_1 = x_1, \ldots, X_n = x_n)$$

$\triangleright$ **model too general to give rise to simple derivations**

Discrete channel models

Properties

For the sake of simplicity, assumptions are made about the model.

Property 1 (Causal channel)**.** *A channel is causal if:*

$$P(Y_1 = y_1, \ldots, Y_m = y_m | X_1 = x_1, \ldots, X_n = x_n)$$
$$= P(Y_1 = y_1, \ldots, Y_m = y_m | X_1 = x_1, \ldots, X_m = x_m)$$

for all m and n such that $m \leq n$.

Consequence. By summing both equality members with respect to $Y_1, \ldots, Y_{m-1}$, we check:

$$P(Y_m = y_m | X_1 = x_1, \ldots, X_n = x_n) = P(Y_m = y_m | X_1 = x_1, \ldots, X_m = x_m)$$

$\longrightarrow$ any output is independent of future inputs

Discrete channel models

Properties

Property 2 (Memoryless causal channel). *A channel is said to be memoryless if, for all $k \geq 2$, we have:*

$$P(Y_k = y_k | X_1 = x_1, \ldots, X_k = x_k, Y_1 = y_1, \ldots, Y_{k-1} = y_{k-1}) \\ = P(Y_k = y_k | X_k = x_k).$$

Consequence. The conditional law governing the channel behavior is entirely determined by the instantaneous conditional laws:

$$P(Y_1 = y_1, \ldots, Y_m = y_m | X_1 = x_1, \ldots, X_n = x_n) = \prod_{k=1}^{m} P(Y_k = y_k | X_k = x_k).$$

$\longrightarrow P(Y_k = y_k | X_k = x_k)$ may be time-dependent.

Discrete channel models

Properties

Noticing that $P(Y_k = y_k|X_k = x_k)$ may depend on the time instant k, we introduce the following property:

Property 3 (Stationary memoryless channel)**.** *A memoryless channel is stationary if, for all $k \geq 1$, we have:*

$$P(Y_k = y_k|X_k = x_k) = P(Y = y_k|X = x_k).$$

Notation. We denote by $(\mathcal{X}, \mathcal{Y}, \Pi)$ any discrete memoryless channel, where Π is the transition matrix defined as:

$$\Pi(i, j) = P(Y = y_j|X = x_i)$$

Discrete channel models

Symmetric channel

A channel is *symmetric* if the rows of its transition matrix all have the same entries up to a permutation, as well as its columns.

Examples. The following transition matrices correspond to symmetric channels.

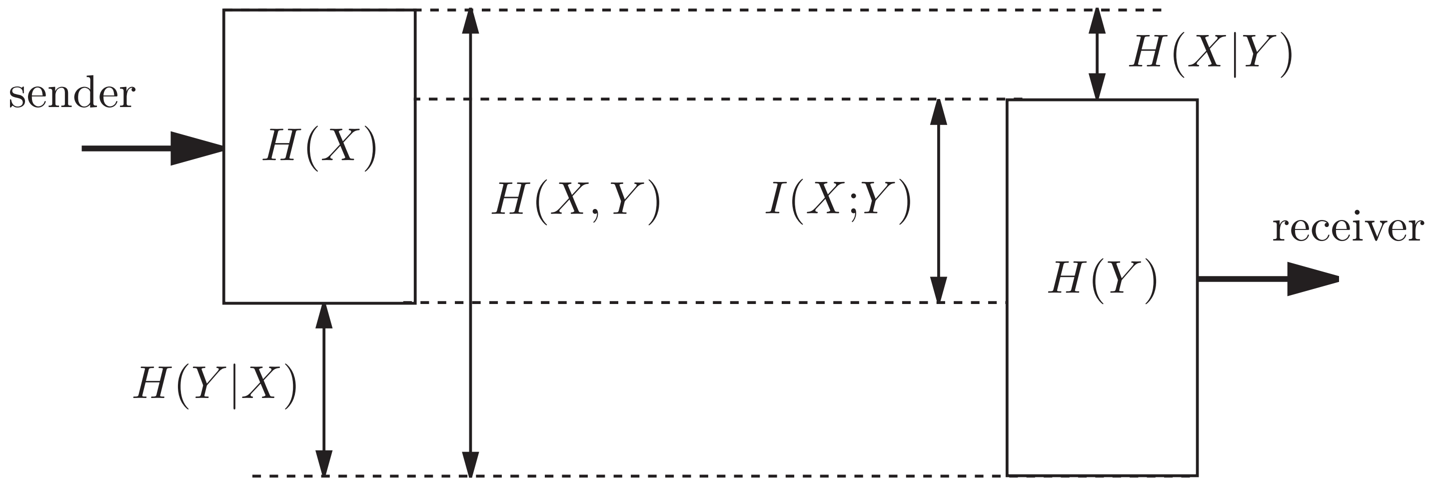
$$\Pi = \begin{pmatrix} p & q & 1-p-q \\ q & 1-p-q & p \\ 1-p-q & p & q \end{pmatrix},$$

$$\Pi = \begin{pmatrix} p & 1-p-q & q \\ q & 1-p-q & p \end{pmatrix},$$

with p and q in $[0,1]$.

Memoryless channel capacity

Introduction



$H(X)$ is the amount of information transmitted through a noiseless channel

$H(X|Y)$ is the amount information required to determine the entry

$I(X;Y)$ is the amount of information transmitted through the channel.

Memoryless channel capacity

Definition

Definition 1. *The information capacity per symbol of a channel is defined as:*

$$C \triangleq \max_{P(X=x)} I(X;Y).$$

Caution. We check that $I(X,Y)$ is a concave function of the law of X. Indeed, by writing $f(x) = -x \log x$, we note that it is a sum of concave functions:

$$\begin{aligned} I(X;Y) &= \sum_i \sum_j p(i,j) \log \frac{p(i,j)}{p(i)\,p(j)} \\ &= \sum_i \sum_j p_i\, p_i(j) \log \frac{p_i(j)}{\sum_i p_i\, p_i(j)} \\ &= \sum_i p_i \left(\sum_j p_i(j) \log p_i(j) \right) + \sum_j f\left(\sum_i p_i p_i(j) \right). \end{aligned}$$

Memoryless channel capacity

Capacity calculation

In the general case, calculating the capacity of a channel is complicated. However, in the case of a symmetric channel, the calculation is easy.

Theorem 1. *The capacity of a symmetric channel* $(\mathcal{X}, \mathcal{Y}, \Pi)$ *is equal to* $I(X;Y)$ *in the case where* X *is governed by a uniform law.*

Proof. The entropy $H(Y|X = x_i) = -\sum_j p_i(j) \log p_i(j)$ is independent of i since the rows i of Π all have the same entries. As a consequence, $H(Y|X)$ is independent of the law of X.

It is easy to check that Y is governed by a uniform law if X is. Indeed,

$$p_j = \sum_i p_i \, p_i(j) = \frac{1}{q} \sum_i p_i(j)$$

is independent of j since the columns of Π all have the same entries. □

Channel capacity calculation

Examples

Noiseless binary channel. The channel outputs are identical to the channel inputs. As a consequence, we have $I(X;Y) = H(X)$ because $H(X|Y) = 0$.

$$C = 1 \text{ Sh/symb}$$

Disfunctional binary channel. This channel always reproduces the same output regardless of the input. Consequently, mutual information $I(X;Y)$ is zero because $H(Y) = H(Y|X) = 0$.

$$C = 0 \text{ Sh/symb}$$

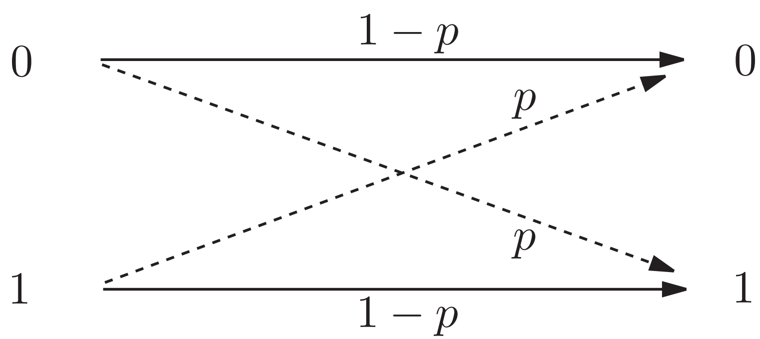
Channel capacity calculation

Binary symmetric channel

The transition matrix of a symmetric binary channel is given by

$$\Pi = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$$

which is represented schematically as follows:



Channel capacity calculation

Binary symmetric channel

In order to evaluate the information capacity of this channel, let us calculate first the mutual information $I(X;Y)$:

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &= H(Y) - P(X=0)H(Y|X=0) - P(X=1)H(Y|X=1). \end{aligned}$$

A simple calculation shows that $H(Y|X=x) = H_2(p)$, where $x \in \{0,1\}$, which leads to:

$$I(X;Y) = H(Y) - H_2(p) \leq \log 2 - H_2(p).$$

As a consequence we have:

$$C = 1 - H_2(p) \text{ Sh/symb}$$

Channel capacity calculation

Binary symmetric channel