

# Bases de la théorie de l'information

235711131723

13 octobre 2022

## 1 Avertissement

Ceci n'est qu'un résumé foireux rédigé par un étudiant sans aucune expérience dans le domaine.

À prendre avec des pincettes.

## 2 Bases de la théorie de l'information

On considère une *source discrète sans mémoire*.

La sortie d'une telle source est une séquence de lettres tirées dans un alphabet  $a_1, \dots, a_K$ .

Chaque lettre de la séquence est statistiquement indépendante des autres et obtenue d'après une loi de probabilité  $P(a_1), \dots, P(a_K)$ .

Pour toute lettre  $a_k$ ,  $1 \leq k \leq K$ , de la source,  $P(a_k)$  est la probabilité pour que cette lettre soit produite, nous aurons donc  $\sum_{k=1}^K P(a_k) = 1$ .

Avant de trouver un moyen de mesurer quantitativement l'information, nous allons essayer de préciser le concept d'information.

Nous l'avons vu, la façon la plus appropriée de décrire un système de communication est d'en donner un modèle probabiliste. Qualitativement, fournir une information consiste à lever une partie de l'incertitude sur l'issue d'une expérience aléatoire. Par exemple, l'observation d'un '0' à la sortie d'un canal binaire bruité augmente la probabilité qu'un '0' ait été émis et diminue la probabilité qu'un '1' ait été émis.<sup>1</sup>

### 2.1 Contenu en information

**Définition.** Soit  $h$  la fonction représentant une mesure du contenu en information de la lettre  $a_k$ . Alors :

$$h(a_k) = -\log_2(p_k)$$

avec  $p_k$  la probabilité d'apparition de la lettre  $a_k$ .

---

1. Source : <https://www.rocq.inria.fr/secret/Nicolas.Sendrier/thinfo.pdf> - p. 12

## 2.2 Entropie d'une source discrète sans mémoire

Il apparaît qu'il existe un lien entre l'information fournie par une source et la distribution de probabilité de la sortie de cette source. Plus l'événement donné par la source est probable, moins la quantité d'information correspondante est grande. Plus précisément, si une lettre  $a_k$  a une probabilité  $P(a_k)$  d'être produite par la source, son information propre sera  $I(a_k) = -\log_2(P(a_k))$ .

Cette définition paraît conforme à l'idée intuitive que l'on peut se faire de l'information, et en particulier on a  $I(a_k) = 0$  si  $P(a_k) = 1$ , c'est-à-dire que l'occurrence d'un événement certain ne peut fournir aucune information. La valeur moyenne de l'information propre calculée sur l'ensemble de l'alphabet revêt une grande importance. Elle est appelée entropie de la source et vaut :

$$H = \sum_{k=1}^K -P(a_k) \log_2(P(a_k))$$

L'entropie d'une source est le nombre moyen minimal de symboles binaires par lettre nécessaires pour représenter la source. Par exemple, si un alphabet contient  $2^L$  lettres équiprobables, il est immédiat que l'entropie de la source correspondante vaut  $L$ . Or il est bien clair que pour représenter  $2^L$  lettres distinctes,  $L$  symboles binaires sont nécessaires. L'entropie d'une source est donnée en bits par seconde et est notée  $H$ .<sup>2</sup>

## 2.3 Propriétés de l'entropie

**Théorème 42.** Soit  $X$  un espace probabilisé discret de cardinal  $K$ . Alors son entropie  $H(X)$  vérifie :

$$H \leq \log_2(K)$$

avec égalité si et seulement la loi de probabilité de  $X$  est uniforme.<sup>3</sup>

## 2.4 Codage de l'information

---

2. Source : <https://www.rocq.inria.fr/secret/Nicolas.Sendrier/thinfo.pdf> - p. 5

3. Source : <https://www.rocq.inria.fr/secret/Nicolas.Sendrier/thinfo.pdf> - p. 15