Introduction à la Théorie de l'information

Raphael Ducatez

Valbonne 2021

Menu

Outline

1. L'information
   - Introduction
   - Le codage.
   - Quelques applications

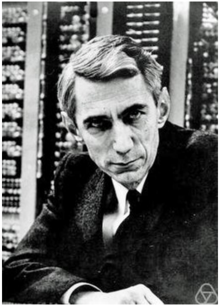
2. L'entropie de Shannon.
   - L'entropie de Shannon.
   - Compression d'information.
   - Correcteur d'erreur.

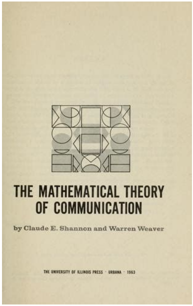
Quelques questions

- Qu'est ce que l'information ?
- Peut-on la mesurer ?
- Comment coder l'information ?
- Comment communiquer (efficacement) ?
- Et si on se trompe ?
- ...

Qu'est ce que l'information ?

Une approche mathématique de l'information





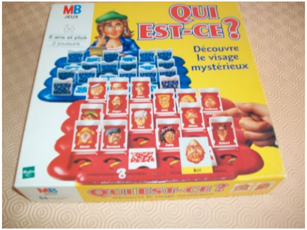
Claude Shannon, (1948).

- Télégraphe (1837).
- Radio (1888).
- Ordinateur (1937)

Outline

1. L'information
   - Introduction
   - Le codage.
   - Quelques applications
2. L'entropie de Shannon.
   - L'entropie de Shannon.
   - Compression d'information.
   - Correcteur d'erreur.

Coder l'information

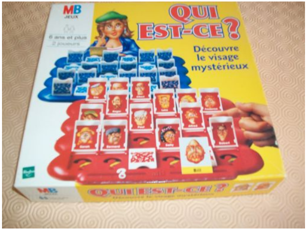


Règles du jeu « Qui est-ce ? » : Trouver le personnage en ne posant que des questions oui ou non.
Stratégie standart : Essayer de faire une dichotomie $2^l \leq N < 2^{l+1}$.

Definition (1)

L'information c'est le nombre de questions « oui ou non » pour caractériser l'objet.

Coder l'information



Règles du jeu « Qui est-ce ? » : Trouver le personnage en ne posant que des questions oui ou non.
Stratégie standart : Essayer de faire une dichotomie.

Definition (1 bis)

L'information de un objet parmi N est $I = \log_2(N)$.

Rappel sur le logarithme

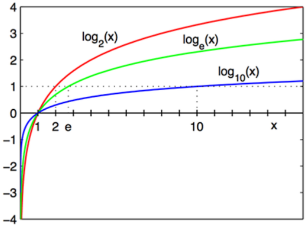
Logarithme en base 2

$$\log_2(2^x) = x$$

Exemple : $\log_2(8) = ?$, $\log_2(\frac{1}{4}) = ?$, $\log_2(1) = ?$, $\log_2(10) = ?$.

Propriété

$$\log_2(a \times b) = \log_2(a) + \log_2(b).$$



Rappel sur le logarithme

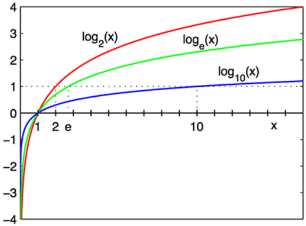
Logarithme en base 2

$$\log_2(2^x) = x$$

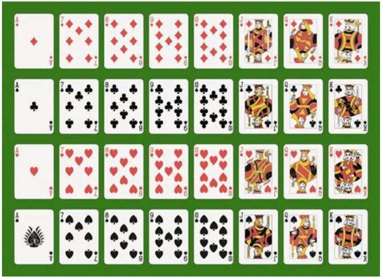
Exemple : $\log_2(8) = 3$, $\log_2(\frac{1}{4}) = -2$, $\log_2(1) = 0$, $\log_2(10) \approx 3,32$.

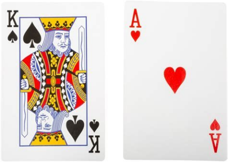
Propriété

$$\log_2(a \times b) = \log_2(a) + \log_2(b).$$



Coder l'information



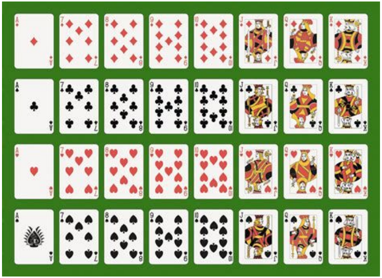


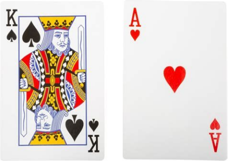
Exemple : la carte est

- $I(«\text{coeur}») = ?$,
- $I(«\text{roi}») = ?$,
- $I(«\text{roi de coeur}») = ?$

.

Coder l'information





Exemple : la carte est

- $I(«\text{coeur}») = 2 = \log_2(4)$,
- $I(«\text{roi}» \ ) = 3 = \log_2(8)$,
- $I(«\text{roi de coeur}») = 5$

Remarque : $I(«\text{roi coeur}») = I(«\text{roi} \ ») + I(«\text{coeur} \ »)$.

Remarque : $I(2 \text{ cartes}) = \log_2(32 \times 32) = 10 = I(1\text{ère carte} \ ) + I(2\text{ème carte})$.

Astuce : Faire des petits groupes.

Exemple : coder un message d'une lettre $\{a, b, c\}$ et d'un chiffre $\{0, 1, 2, 3, 4\}$.
Codage 1

$$\begin{cases} a \leftrightarrow 00 \\ b \leftrightarrow 01 \\ c \leftrightarrow 10 \end{cases} \quad \begin{cases} 0 \leftrightarrow 000 \\ 1 \leftrightarrow 001 \\ 2 \leftrightarrow 010 \\ 3 \leftrightarrow 011 \\ 4 \leftrightarrow 100 \end{cases}$$

par exemple $b2 \leftrightarrow 01\,010$. (longueur 5)
Codage 2 ?

Astuce : Faire des petits groupes.

Coder un message d'une lettre $\{a, b, c\}$ et d'un chiffre $\{0, 1, 2, 3, 4\}$.
Codage 1

$$\begin{cases} a \leftrightarrow 00 \\ b \leftrightarrow 01 \\ c \leftrightarrow 10 \end{cases} \quad \begin{cases} 0 \leftrightarrow 000 \\ 1 \leftrightarrow 001 \\ 2 \leftrightarrow 010 \\ 3 \leftrightarrow 011 \\ 4 \leftrightarrow 100 \end{cases}$$

par exemple b2↔01 010.

Codage 2 : 15 message donc $\leq 16 = 2^4$. Seulement de longueur 4 !

$$\begin{cases} a0 \leftrightarrow 0000 \\ a1 \leftrightarrow 0001 \\ \dots \\ c4 \leftrightarrow 1110 \end{cases}$$

$$(\log_2(3) + \log_2(5) \approx 3,9 < 4)$$

Coder avec un autre alphabet ?

Autres alphabet

Avec un alphabet à k lettre, on code $\log_2(k)$ d'information par lettre.

Exemple pour un alphabet avec quatre lettres on code 2 d'information par lettre.

$$0 \leftrightarrow 00$$
$$1 \leftrightarrow 01$$
$$2 \leftrightarrow 10$$
$$3 \leftrightarrow 11$$
$$10323 \leftrightarrow 01\,00\,11\,10\,11$$

Outline

1. L'information
   - Introduction
   - Le codage.
   - Quelques applications
2. L'entropie de Shannon.
   - L'entropie de Shannon.
   - Compression d'information.
   - Correcteur d'erreur.

Le jeu « Code name ».

Règles du jeu : Faire deviner à ses coéquipier le plus de cases de son équipe en ne disans qu'un seul mot..
Ici c'est un codage dont l'alphabet est l'ensemble des mots de la langue française :
1 case : 25 possibilités ($I \approx 4,6$).
2 cases : 300 possibilités ($I \approx 8,2$)
3 cases : 2300 possiblités ($I \approx 11,2$)
4 cases : 12650 possibilités (très très difficile)

L'énigme de la fausse pièce

« Vous disposez de 9 pièces : 8 sont en or et sont identiques, une est une fausse pièce peinte en doré mais est plus légère que les autres. Pouvez déterminer laquelle en seulement 2 pesées ? »

L'énigme de la fausse pièce

« Vous disposez de 9 pièces : 8 sont en or et sont identiques, une est une fausse pièce peinte en doré mais est plus légère que les autres. Pouvez déterminer laquelle en seulement 2 pesées ? »

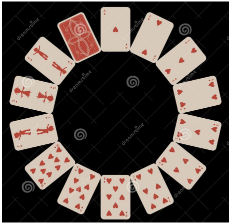
Solution : Peser deux groupes de 3 pièces si le plateau penche d'un coté on sait dans quel groupe de trois pièces est la fausse pièce et si c'est à l'équilibre elle se trouve dans le dernier groupe. Faire la même chose avec les trois pièces séléctionnées.
Information : À chaque pesée $I = \log_2(3)$.
$I(\text{fausse pièce}) = \log_2(9) = 2 \times \log_2(3)$. (donc 2 pesées).

Un petit tour de « magie ».

« Le magicien donne le paquet de cartes à une personne dans la salle, il laisse la personne couper le deck (et replacer les deux parties) puis demande à plusieurs personne de faire de même. Il demande ensuite à cinq personnes de piocher chacun une carte du deck. Ensuite le magicien essaie de lire les pensées de ces cinq personnes et leur demande de ce concentrer sur les cartes qu'ils ont. Pour l'aider contre le bruit qui vient des autres personnes, il demande aux personnes avec une carte noire de faire un pas en avant. Il devine ensuite chacunes des cinq cartes piochées »

Un petit tour de « magie ».

« *Le magicien donne le paquet de cartes à une personne ... couper le deck ... il demande aux personnes avec une carte noire de faire un pas en avant.* »



Inconnue : la coupure : 32 possibilités $I = \log_2(32) = 5$.
Information codé : 5 « rouges ou noires » $= 5$.

Outline

1. L'information
   - Introduction
   - Le codage.
   - Quelques applications

2. L'entropie de Shannon.
   - L'entropie de Shannon.
   - Compression d'information.
   - Correcteur d'erreur.

Information et probabilités

Quelle est l'information de « ce n'est pas un roi de coeur » ?
Quelle est l'information pour le résultat d'un dé pipé ?

Definition (2)

$I(X = a) = -\log_2 \mathbb{P}(X = a)$.

Exemple : Tirer une carte d'un jeu de 32 :

$$I(\text{roi de coeur}) = -\log_2(\frac{1}{32}) = 5.$$

(Tirage indépendant) $\mathbb{P}(X = a, Y = b) = \mathbb{P}(X = a) \times \mathbb{P}(Y = b)$

$$I(X = a, Y = b) = I(X = a) + I(Y = b)$$

Information et probabilités

Quelle est l'information de « ce n'est pas un roi de coeur » ?
Quelle est l'information pour le résultat d'un dé pipé ?

Definition (2)

$I(X = a) = -\log_2 \mathbb{P}(X = a)$.

(Probabilité conditionnelle) $\mathbb{P}(A|B) = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$

$$I(A|B) = I(A\cap B) - I(B).$$

Exemple : I(roi de coeur | rouge) $=$?.

Information et probabilités

Quelle est l'information de « ce n'est pas un roi de coeur » ?
Quelle est l'information pour le résultat d'un dé pipé ?

Definition (2)

$I(X = a) = -\log_2 \mathbb{P}(X = a).$

(Probabilité conditionnelle) $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$

$$I(A|B) = I(A \cap B) - I(B).$$

Exemple : $\mathbb{P}(\text{roi de coeur} \mid \text{rouge}) = \frac{1/32}{1/2} = \frac{1}{16}$.
$I(\text{roi de coeur} \mid \text{rouge}) = 4$.

Codage optimale

Sur un canal, des caractères sont transmis « aléatoirement » selon une certaine loi de probabilité. Trouver la manière de les coder pour minimiser la longueur du signal.

Exemple : $\{a, b, c, d\}$: $\mathbb{P}(a) = \frac{1}{2}$, $\mathbb{P}(b) = \frac{1}{4}$, $\mathbb{P}(c) = \frac{1}{8}$ et $\mathbb{P}(d) = \frac{1}{8}$.

$a \leftrightarrow 00 \qquad c \leftrightarrow 10$

$b \leftrightarrow 01 \qquad d \leftrightarrow 11$

$acbaadba \leftrightarrow 00\,10\,01\,00\,00\,11\,01\,00$

En moyenne 2 par caractère

$a \leftrightarrow 0 \qquad c \leftrightarrow 110$

$b \leftrightarrow 10 \qquad d \leftrightarrow 111$

$acbaadba \leftrightarrow 0\,110\,10\,0\,0\,111\,10\,0$

En moyenne $1,75$ par caractère

« l'entropie de Shannon »

Entropie

C'est « information moyenne » que l'on obtient à chaque lancé $S(X) = -\sum_a \mathbb{P}(X = a) \times \log_2 \mathbb{P}(X = a)$.

Exemple : $\{a, b, c, d\}$: $\mathbb{P}(a) = \frac{1}{2}$, $\mathbb{P}(b) = \frac{1}{4}$, $\mathbb{P}(c) = \frac{1}{8}$ et $\mathbb{P}(d) = \frac{1}{8}$.

$$\begin{aligned} S &= -\frac{1}{2} \times \log_2(\frac{1}{2}) - \frac{1}{4} \times \log_2(\frac{1}{4}) - \frac{1}{8} \times \log_2(\frac{1}{8}) - \frac{1}{8} \times \log_2(\frac{1}{8}) \\ &= 1,75 \end{aligned}$$

Question : Pour une variable aléatoire à N éléments quelle distribution dont l'entropie est la plus grande ?

Théorème

Le plus court codage possible pour un canal est donné par l'entropie de Shannon.

« l'entropie de Shannon »

Entropie

C'est « information moyenne » que l'on obtient à chaque lancé $S(X) = -\sum_a \mathbb{P}(X=a) \times \log_2 \mathbb{P}(X=a)$.

Exemple : $\{a, b, c, d\}$: $\mathbb{P}(a) = \frac{1}{2}$, $\mathbb{P}(b) = \frac{1}{4}$, $\mathbb{P}(c) = \frac{1}{8}$ et $\mathbb{P}(d) = \frac{1}{8}$.

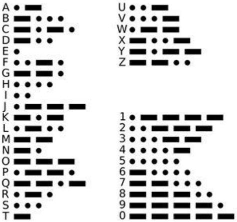
$$
\begin{aligned}
S &= -\frac{1}{2} \times \log_2(\frac{1}{2}) - \frac{1}{4} \times \log_2(\frac{1}{4}) - \frac{1}{8} \times \log_2(\frac{1}{8}) - \frac{1}{8} \times \log_2(\frac{1}{8}) \\
&= 1,75
\end{aligned}
$$

Réponse : La loi uniforme $\mathbb{P}(a) = \ldots = \mathbb{P}(d) = \frac{1}{N}$ (Jensen !).

Théorème

Le plus court codage possible pour un canal est donné par l'entropie de Shannon.

L'entropie de la langue anglaise ?

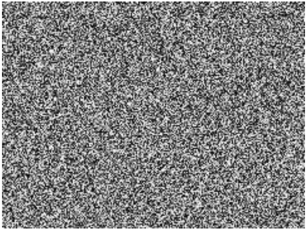


Codage la langue Anglaise (27 lettres).

- $F0 = \log_2(27) \approx 4,76$.
- $F1 = S(1 \text{ lettre}) \approx 4,03$
- $F2 = S(2 \text{ lettre}) \approx 3,32$
- $F3 = S(3 \text{ lettre}) \approx 3,1$
- $F_m = S(\text{ mots}) \approx 2,14$

Outline

Compression d'image



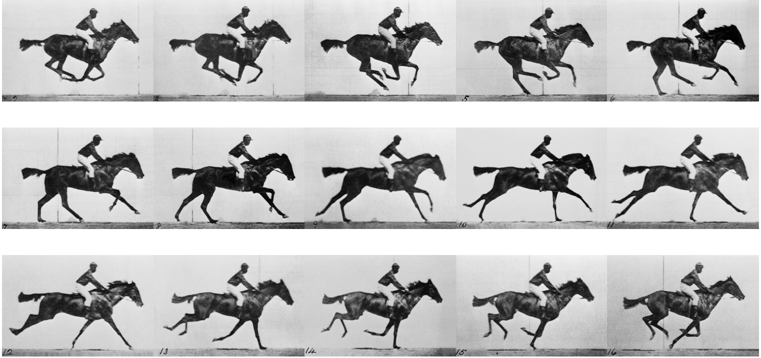
« bruit blanc ».

image « normal ».

$$5700 \times 3800 = 4Mo \ (??).$$

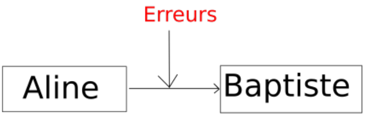
Compression de vidéo

1 image $\approx 1 Mo$.
1 vidéo (film d'une heure 30 min) $\approx 1 Go$.
1 seconde = 24 image $\Rightarrow$ 90 minutes = 129600 images
$\approx 130 Go\,(??)$



Outline

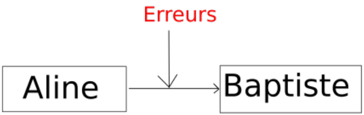
Code correcteur d'erreurs



1. Détecter les erreurs ?
2. Corriger les erreurs ?

(Envoyer un message plus long)

Code correcteur d'erreurs



1. Détecter les erreurs ?
2. Corriger les erreurs ?

Exemple :

1. Ajouter un chiffre = la somme des autres chiffres. $110 \leftrightarrow 1100$
2. Répéter 3 fois la même lettre $110 \leftrightarrow 111\,111\,000$ (pas vraiment optimale...).

Le code de Hamming

Message de longeur 3 "*abc*" de $\{0,1\}$ (exemple 101) . Le coder avec un message de longeur 7 :

$$\begin{cases} x_1 = a & x_5 = a + c \\ x_2 = b & x_6 = b + c \\ x_3 = a + b & x_7 = a + b + c \\ x_4 = c & \end{cases} \begin{cases} x_1 = 1 & x_5 = 1 + 1 = 0 \\ x_2 = 0 & x_6 = 0 + 1 = 1 \\ x_3 = 1 + 0 = 1 & x_7 = 1 + 0 + 1 = 0 \\ x_4 = 1 & \end{cases}$$

$$101 \leftrightarrow 1011010$$

Exercice : Il y a une erreur dans 1000110, quel était le message codé ?

Le code de Hamming

Message de longueur 3 "*abc*" de $\{0,1\}$ (exemple 101) . Le coder avec un message de longueur 7 :

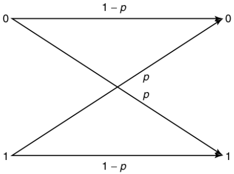
$$\begin{cases} x_1 = a & x_5 = a + c \\ x_2 = b & x_6 = b + c \\ x_3 = a + b & x_7 = a + b + c \\ x_4 = c & \end{cases} \begin{cases} x_1 = 1 & x_5 = 1 + 1 = 0 \\ x_2 = 0 & x_6 = 0 + 1 = 1 \\ x_3 = 1 + 0 = 1 & x_7 = 1 + 0 + 1 = 0 \\ x_4 = 1 & \end{cases}$$

$$101 \leftrightarrow 1011010$$

Exercice : Il y a une erreur dans 1000110, quel était le message codé ?
Réponse : 110

Code correcteur d'erreurs

Sur chaque lettre, il y a une erreur avec probabilité p.

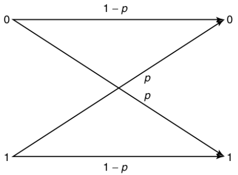


En moyenne, combien d'information Aline peut-elle transmettre si

1. $p = 0$?
2. $p = 1$?
3. $p = \frac{1}{2}$?

Code correcteur d'erreurs

Sur chaque lettre, il y a une erreur avec probabilité p.

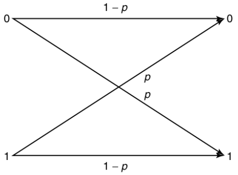


En moyenne, combien d'information Aline peut-elle transmettre si

1. $p = 0$: 1 (il n'y a aucune erreur)
2. $p = 1$: 1 (il suffit de remplacer 0 par 1 et inversement)
3. $p = \frac{1}{2}$: 0 (il est impossible de faire la différence en le 0 et le 1 initial)

Code correcteur d'erreurs

Sur chaque lettre, il y a une erreur avec probabilité p.



l'entropie de l'erreur $S(p) = p\log_2 p + (1-p)\log_2(1-p)$.

Limite donné par Shannon

L'information obtenue pour chaque lettre est $1 - S(p)$.

Concrètement il faut un message de longueur $\frac{L}{1-S(p)}$ pour coder un message de longueur L tout en détectant et corrigeant les erreurs.

Merci de votre attention.