* Codage source *

- Codage source ≡ Numérisation de l'information
  ↳ information analogique → information numérique

- 3 phases ⇒
  - Echantillonnage $T_e = \frac{1}{F_e}$
  - Quantification ≡ mesurer l'échantillon
  - Codage

⇓

Réduire la qté d'information délivrée source (Éliminer la redondance)

* Code à longueur fixe :
- représenté à la m̂ nbre de bits
  $S = \{10, 01, 11, 00\}$
  ↳ de longueur fixe = 2

à longueur variable :
- représenté à différent de bits
  $S = \{101, 100, 10\}$
  ou $S = \{10, 111, 1\}$

* Code préfixe
$S = \{01, 110, 001\}$
$S = \{111, 110, 101\}$

non préfixe
$S = \{001, 110, 11\}$
$S = \{011, 00, 11, 110\}$

$S = \{100, 01, 10, 001, 110\}$ non préfixe

* La longueur d'un mot code :
$$\bar{n} = \sum_{i=1}^{M} p_i n_i$$

$S = \{S_1, S_2, \ldots, S_M\}$ avec n_i : Longueur de S_i (nbre de bit de S_i)

- M : nbre des mots code ;
- p_i = prob. de S_i

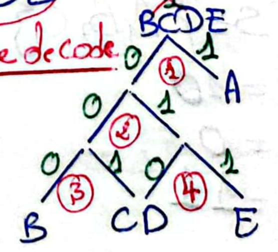
* 1 * Code de Huffman
$S = \{A, B, C, D, E\}$

$p_i = \left\{\frac{15}{39}, \frac{7}{39}, \frac{6}{39}, \frac{6}{39}, \frac{5}{39}\right\}$

1 A $\frac{15}{39}$	1 A $\frac{15}{39}$	1 A $\frac{15}{39}$	1 A $\frac{15}{39}$	1 A $\frac{15}{39}$
0 B $\frac{7}{39}$	0 B $\frac{7}{39}$	1 DE $\frac{11}{39}$	0 DE $\frac{11}{39}$	0 BC $\frac{13}{39}$
1 C $\frac{6}{39}$	1 C $\frac{6}{39}$	0 B $\frac{7}{39}$	0 BC $\frac{13}{39}$	1 DE $\frac{11}{39}$
0 D $\frac{6}{39}$	1 DE $\frac{11}{39}$	1 C $\frac{6}{39}$		
1 E $\frac{5}{39}$				

1 A $\frac{15}{39}$	1 A	0 BCDE $\frac{24}{39}$	BCDEA
0 BC	0 BCDE $\frac{24}{39}$	1 A $\frac{15}{39}$	
1 DE			

Arbre de code



A = 1
B = 000
C = 001
D = 010
E = 011

$\bullet\ \bar{n} = \sum_{i=1}^{5} n_i p_i = 1\times\frac{15}{39} + 3\times\frac{7}{39} + 3\times\frac{6}{39} + 3\times\frac{6}{39} + 3\times\frac{5}{39}$

$\Rightarrow \bar{n} = 2{,}23$ bits

$\bullet\ H(S) = -\sum_{i=1}^{5} p_i \log_2(p_i) = \frac{-1}{39}\left[15\log_2\left(\frac{15}{39}\right) + 7\log_2\left(\frac{7}{39}\right) + 12\log_2\left(\frac{6}{39}\right) + 5\log_2\left(\frac{5}{39}\right)\right] = 2{,}18\,\text{sh}$

⊛ Taux d'émission de la source : $T_{émission} = D_s\, H(S)$

(bits/s) — D_s : (Débit) cadense (symbole/s) ; $H(S)$: Entropie (bits/symb)

⊛ Redondance : $R_e = 1 - \frac{H(S)}{H_{max}(S)}$

avec $H_{max} = \log_2(A)$ — A : nbre des variable.

* La capacité du canal binaire : $C = D_b \log_2(n)$

le débit du canal binaire — (cas d'un canal binaire)

* Cond : Si $T_{émission} < C \Rightarrow$ on peut transmettre la totalité de l'information

* 2* Codage de Shanon - Fano : $P = \left\{\frac{1}{4}, \frac{5}{8}, \frac{1}{8}\right\}$

$S = \{A, B, C\}$

A $\frac{1}{4}$	B $\frac{5}{8}$	0		$A = 10$
B $\frac{5}{8}$	$\longrightarrow$ A $\frac{1}{4}$		0	$\longrightarrow$ $B = 0$
C $\frac{1}{8}$	C $\frac{1}{8}$	1	1	$C = 11$

$\bar{n} = \sum_{i=1}^{3} n_i p_i = \frac{1}{4}\times 2 + 1\times\frac{5}{8} + \frac{1}{8}\times 2 = 1{,}37$ bits

* Efficacité du code : $\text{Efficacité} = \frac{H(S)}{\bar{n}}$

%

$\Rightarrow$ Tjrs Huffmann meilleur que Shannon-Fano

$D_s = \frac{1}{\sum_{i=1}^{n} p_i t_i}$ débit moyen

symb/s

Théorie de l'information

* **Alphabet** : $S = \{S_1, S_2, \dots, S_N\}$

* **Source Aléatoire** : $P(S = s_N) = p_N$, $k = 1, \dots, N$ avec $\sum_{k=1}^{N} p_k = 1$

* **La quantité d'information** :

$I(S_k) = -\log_2 p_k$

- $p_k = \frac{1}{2} \Rightarrow I(S_k) = 1$ bit
- $p_k = 1 \Rightarrow I(S_k) = 0$ bit
- $0 \leq p_k \leq 1 \Rightarrow I(S_k) = 0$ bit
- $I(S_k) > I(S_i)$ pour $p_k < p_i$
- Pour S_k et S_i sont statiquement indép : $I(S_k S_i) = I(S_k) + I(S_i)$

* **Source d'information** :
→ symbole S_1, Probabilité P_1
S_2 — P_2
⋮
→ symbole S_N, Probabilité P_N

- $S = \{S_1, \dots, S_N\}$
- $P(S_k) = \frac{1}{N}$, $k = 1 \dots N$

* **Entropie de la source S** : $H(S) = -\sum_{k=1}^{N} p_k \log_2(p_k)$

- Si N symboles sont équiprobable ($P(S_k) = \frac{1}{N}$, $k = 1, \dots, N$)

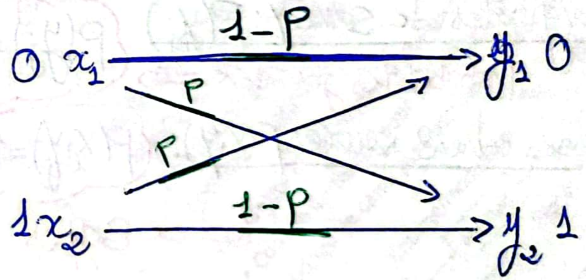
Alors $H(S) = \log_2(N)$

- Si les symbole ne sont pas équiprobable et $H(S) = 0{,}88$ Sh
- Maintenant les symboles sont équiprobable et $H(S) = 1$ Sh

⇒ plus l'entropie est grande quand les messages sont équiprobable.

* **Entropie Conditionnelle** :

Source d'entrée $X(x_1, x_2)$ — source de sortie $Y(y_1, y_2)$



x_1 et x_2 sont équiprobables : $p(x_1) = p(x_2) = \frac{1}{2}$

- **Quantité d'information mutuelle $I(X,Y)$** :

$I(X,Y) = H(Y) - H(Y/X)$ ②

ou

$I(X,Y) = H(X) + H(Y) - H(X,Y)$ ①

- **Entropie jointe $H(X,Y)$** :

① $H(X,Y) = -\sum_i \sum_j p(x_i, y_j) \log_2(P(x_i, y_j))$

• Entropie conditionnelle moyenne $H(Y/X)$

$$(2)\quad H(Y/X) = \sum_i p(x_i)\, H(Y/x_i)\ (3)$$

$$\text{avec } (3)\ H(Y/x_i) = -\sum_i p(Y/x_i) \log_2 \left(p(Y/x_i)\right)$$

• Quantité d'information d'un caractère $I(x_j)$:

$$I(x_j) = -\log_2 (P(x_j))$$

• Proba. conditionnelle $P(X/Y)$: $P(X/Y) = p(X,Y) / P(Y)$

Loi des Bayes

• Proba. marginales de source Y : $P(y_j)$: $P(y_j) = \sum_{i=1}^{N} p(x_i, y_j)$

Probabilités jointes $\Longrightarrow$

• Proba. jointes entre 2 sources $P(X,Y)$: $P(x_i, y_j) = P(y_j/x_i)\, P(x_i)$

Loi de Bayes

Dans ce cas :

• Proba. cond. : $\begin{cases} P(y_1/x_1) = 1-p \\ P(y_2/x_2) = 1-p \end{cases}$ $\begin{cases} p(y_2/x_1) = p \\ p(y_1/x_2) = p \end{cases}$

• Proba. jointes entre 2 sources : $\boxed{P(X,Y)}$

• $P(x_1, y_1) = P(y_1/x_1) P(x_1) = \frac{1-p}{2}$

• $P(x_2, y_2) = P(y_2/x_2) p(x_2) = \frac{p}{2}$

• $P(x_2, y_2) = \frac{1-p}{2}$ et • $P(x_2, y_1) = \frac{p}{2}$

• Proba. marginales de Y :

• $P(y_1) = \sum_{i=1}^{2} P(x_i, y_1) = P(x_1, y_1) + p(x_2, y_2) = \frac{1}{2}$

• $P(y_2) = \frac{1}{2}$

• Proba. cond. $P(X/Y)$: (Bayes)

• $P(x_1/y_1) = \dfrac{P(x_1, y_1)}{P(y_1)} = \dfrac{\frac{1}{2}(1-p)}{\frac{1}{2}} = 1-p$

• $P(x_1/y_2) = p$ • $P(x_2/y_1) = p$ • $P(x_2/y_2) = 1-p$

• Quantité d'inf. d'un caract. : $I(x_j)$

$I(x_1) = -\log_2 (P(\frac{1}{2})) = \log_2 2 = 1\,bit = I(x_2) = I(y_1) = I(y_2)$

• Entropie de chaque source : $H(X) = H(Y) = 1\,bit/symbole = 1\,Sh$

$$\left[H(X) = \log_2 (\underset{=2}{K}) \right]$$

• Entropie cond. moyenne : $H(Y/X) = p(x_1)\left[-P(y_1/x_1)\log_2(P(y_1/x_1)) - P(y_2/x_1)\log_2(P(y_2/x_1))\right] + p(x_2)\left[-P(y_1/x_2)\log_2(P(y_1/x_2)) - P(y_2/x_2)\log_2(P(y_2/x_2))\right]$

$= -(1-p)\log_2(1-p) - p\log_2(p)$

• $H(X/Y) = H(Y/X)$ car canal binaire symétrique

• Entropie jointe $H(X,Y)$: $H(X,Y) = 1 - p\log_2(p) - (1-p)\log_2(1-p)$

• ~~Entropie~~ Qté d'inf. mutuelle $I(X,Y)$:

$$I(X,Y) = 1 + (1-p)\log_2(1-p) + p\log_2(p)$$