

# TD : Th. de l'informat.

1.

-  $\log_2 p_j$  = la qté d'info associé à  $p_j$

Entropie = qté moyenne d'info

val de

(N fini)

Exo 1: Max d'entropie pour loi discrète à support fini : loi  $P : \{p_0, p_1, \dots, p_{N-1}\}$  N états

entropie :  $H(X) = -\sum_{j=0}^{N-1} p_j \log_2 p_j$  unité : bit

1)  $\forall q \ H(P) \geq 0$  :  $p_j \in [0, 1]$  donc  $\log_2 p_j \leq 0$  donc  $-p_j \log_2 p_j \geq 0$

$\sum p_j \log_2 p_j \rightarrow 0$  si  $p = 0$   
ou  $p = 1$  et comme les probas normalisées  $\sum_{j=0}^{N-1} p_j = 1$

Donc seul cas pour avoir  $H(P)=0$  :  $\exists i \in [0, N-1]$ ,  $p_i=1$  et  $\forall j \in [0, N-1] \neq i$  :  $p_j=0$

2)  $\forall q \ \sum p_j \log_2 \frac{p_j}{q_j} \geq 0$

$\forall x \in \mathbb{R}^*$ :

$\forall x \log x \leq x-1$  ( $=$  si  $x=1$ )

$$\log \frac{p_j}{q_j} \leq \frac{p_j}{q_j} - 1$$

$$\sum_{j=0}^{N-1} p_j \log \frac{p_j}{q_j} \leq \underbrace{\sum_{j=0}^{N-1} p_j}_{1} - \underbrace{\sum_{j=0}^{N-1} p_j}_{1}$$

$$\sum_{j=0}^{N-1} p_j \log_2 \frac{p_j}{q_j} > 0$$

} Egalité qdssi  $\forall j \in [0, N-1] : q_j = p_j$   
ie  $P = Q$

\*  $D(P||Q) = \sum_{j=0}^{N-1} p_j \log_2 \frac{p_j}{q_j} \geq 0$  (divergence de Kullback) (parfois appelée distance)

$$x \ H(P) = -\sum_{j=0}^{N-1} p_j \log_2 p_j \leq -\sum_{j=0}^{N-1} p_j \log_2 q_j \quad (*)$$

3) et 4)  $\forall q \ H(P) \leq \log_2 N$

(\*) en choisissant la loi  $Q$  : loi uniforme :  $\forall j \in [0, N-1] : q_j = \frac{1}{N}$

$$H(P) \leq -\sum_{j=0}^{N-1} p_j \log_2 \frac{1}{N}$$

$$\leq \log_2 N \sum_{j=0}^{N-1} p_j$$

$H(P) \leq \log_2 N$  égalité qdssi  $P = Q$  ie  $P$  = loi uniforme à  $N$  états.

Exo 2 : Octet d'information

Couple de va  $X_1, X_2 \quad X_1, X_2 \in \{0, 1\}$  :  $(X_1, X_2) \in \{(0,0), (0,1), (1,0), (1,1)\}$   
4 états possibles pour le couple.

1)  $H$  max si les 4 états sont équiprobables (cf q. précédente)

Donc si  $(X_1, X_2)$  loi uniforme à 4 états.

$H_{\max}^{(X_1, X_2)} = \log_2 4 = 2$  bits d'information.

2) Si  $(X_1, X_2) \sim$  uniforme à  $N=4$  états. Loi de  $X_1$ ? Loi de  $X_2$ ?

$X_2$	0	1	$P(X_1)$
$X_1$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
1	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
$P(X_2)$	$\frac{1}{2}$	$\frac{1}{2}$	

$$P(X_1=i) = \sum_{j=0}^{N-1} P(X_1=i, X_2=j)$$

$$P(X_1=1) = P(X_1=1, X_2=0) + P(X_1=1, X_2=1)$$

Donc  $X_1$  et  $X_2$  uniforme à 2 états

3)  $X_1 \subset \mathcal{U}(2)$  ?  $\Rightarrow (X_1, X_2) \subset \mathcal{U}(4)$   
 $X_2 \subset \mathcal{U}(2)$   
NON du le cas gal

Ctr-ex:

$X_1$	0	1	$P(X_1)$
0	1/2	0	1/2
1	0	1/2	1/2
$P(X_2)$	1/2	1/2	

pas uniforme

Rq:  $X_1, X_2$  non indépendants car  $p(X_1=0 \wedge X_2=0) = \frac{1}{2} \neq p(X_1=0) \cdot p(X_2=0) = 0$

4)  $(X_1, X_2) \subset \mathcal{U}(4)$   $\Rightarrow X_1, X_2$  ind.

En effet:  $\forall i, \forall j : P(X_1=i, X_2=j) = P(X_1=i) \cdot P(X_2=j)$

5) Conclusion:  $\begin{cases} X_1 \subset \mathcal{U}(2) \\ X_2 \subset \mathcal{U}(2) \\ X_1, X_2 \text{ indépendant} \end{cases}$

6) Gal: actes:  $N = 2^8$  états Indépendance mutuelle

Exo 3. Max d'entropie ss contrainte de support fini mais contrainte de moyenne.

loi Q:  $\{q_k, k \in \mathbb{N}\}$

moyenne:  $\mu = \sum_{k=0}^{+\infty} k \cdot q_k$  (espérance)

1)  $H(Q_{geo})$   $Q_{geo}: q_k = \alpha \beta^k$  avec  $0 < \beta < 1, \alpha = \frac{1}{1+\mu}, \beta = \frac{\mu}{1+\mu}$

$H(Q_{geo})$  en fonction de  $\mu$ ?

$$H(Q_{geo}) = - \sum_{k=0}^{+\infty} q_k \log_2(q_k) = - \sum_{k=0}^{+\infty} q_k (\underbrace{k \cdot \log_2 \beta + \log_2 \alpha}_{\text{on utilise } \log_2 \alpha = \log_2(1+\mu)}) = - \log_2 \underbrace{\sum_{k=0}^{+\infty} q_k}_{1} - \log_2 \underbrace{\sum_{k=0}^{+\infty} k q_k}_{\mu}$$

$$H(Q_{geo}) = - \log_2 \alpha - \mu \log_2 \beta = - \log_2 \left( \frac{1}{1+\mu} \right) - \mu \log_2 \left( \frac{\mu}{1+\mu} \right) = \frac{1}{1+\mu} \log_2(1+\mu) - \mu \log_2 \mu \quad (**)$$

2) Rq HP de moy  $\mu$ :  $H(P) \leq H(Q_{geo})$

(\*) avec Q loi geo  $H(P) \leq - \sum_{j \geq 0} p_j \log_2(p_j)$

on réutilise les m<sup>e</sup> calculs ?

$$\leq (1+\mu) \cdot \log_2(1+\mu) - \mu \log_2 \mu$$

= si P géométrique

3) ex:  $\mu = 1$

Uniforme  $\approx N = 3$  états  $\begin{matrix} 0 & 1 & 2 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{matrix}$

$$H(\text{Unif}) = \log_2(3)$$

$$H(\text{geo}) = (1+1) \log_2(1+1) - 1 \log_2(1)$$

$$= 2 = \log_2(4) > \log_2(3)$$

# TD : Th. de l'info.

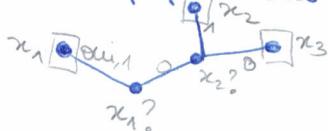
## 2. L'entropie de Shannon en tant que borne

Exo 1 : Trouver l'entropie, version Roberval

1 objet à identifier parmi  $A_x = \{x_1, x_2, \dots, x_N\}$  ddp  $P_x = p_1, p_2, \dots, p_N$   
en posant des q. à réponse à  $D$  états

ex :  $N = 3$  objets

$D = 2$  rep possibles : Oui (1) / Non (0)



$$x_1 \rightarrow 1 : l_1 = 1 \quad \text{nb quest}^*$$

$$x_2 \rightarrow 01 : l_2 = 2$$

$$x_3 \rightarrow 00 : l_3 = 2$$

$$\text{nb moyen de questions : } \bar{l} = \sum_{i=1}^{N \text{ (mots)}} p_i l_i$$

$$(*) l_{\max} = \bar{l} \geq \frac{H(x)}{\log_2(D)} \stackrel{\text{def}}{=} H_D(x)$$

$$l_{\max} = \max_{i=1, \dots, N} \{l_i\}$$

Rg : Si  $X \sim$  loi uniforme à  $N$  états  
et  $l_i = l \forall i = 1, \dots, N$   $\bar{l} = l = l_{\max}$   
 $H(x) = \log_2(N)$   $(*) \Leftrightarrow \log_2(N) \leq l$

$N \leq D^l$  → en base  $D$  avec  $l$  elt on peut pas coder  
+ de  $N$  "mots"

Ex : En base 2, long 3 → on code 8 elt.

Questions générales :

Q1. Qté max d'info en 1 pesée ?

$H_{\max}(1 \text{ pesée}) \leq \log_2(3)$  → 3 états possibles : → penche à d.  
→ eq.  
→ penche à g.  
ssi les 3 états sont equiprobables

$$H_{\max}(1 \text{ pesée}) = \log_2(3) \approx 1,58 \text{ bits d'info en 1 pesée}$$

Q2. Choix de la 1<sup>e</sup> pesée pour avoir  $H_{\max}(1 \text{ pesée})$  ?

Il n'y a pas de choix de la 1<sup>e</sup> pesée pour avoir  $H_{\max}(1 \text{ pesée})$ .  
Le résultat dépend de la 1<sup>e</sup> pesée.

étais	✓	↓	↙
$m=1$	$1/3$	$2/3$	$1/3$
$m=2$	$2/3$	$1/3$	$2/3$
$m=3$	$3/3$	$3/3$	$3/3$
$m=4$	$4/3$	$1/3$	$4/3$

loi uniforme .

Résolution de Q1 : Trouver l'entropie sachant que ~~on~~ + lourd ?

1  $H(Q_1) = \log_2(9)$  : il y a 9 réponses (objets) possibles, tous équiprobables.

Q1 : VA d'alphabet  $A_{Q_1} = \{1, 2, 3, \dots, 8, 9\}$   $P_{Q_1} = \frac{1}{9}$

$$\text{d'où } H(Q_1) = \log_2(\text{Nbr d'états}) = \log_2(9) = 2 \cdot \log_2(3)$$

$$H(Q_1) = \log_2(9) = \log_2(3^2) = 2 \log_2(3) = 2 H(1 \text{ pesée}) \approx 3,16 \text{ bits}$$

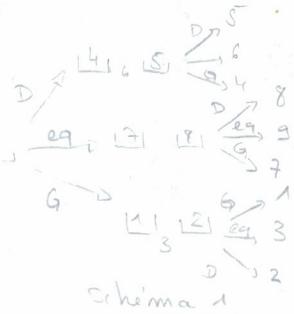
2 - possible de trouver 1 pesée en 1 pesée ( $\Rightarrow l_{\max} = 1$ ) ?

NON car  $H(Q_1) > H(1 \text{ pesée})$

Par imp en  $l_{\max} = 2$  pesées puisqu'en 2 pesées on peut espérer :  $2H_{\max}(1 \text{ pesée}) = H(Q_1)$

$$\text{ou en utilisant } (*) \quad l_{\max} \geq \frac{H(Q_1)}{\log_2(3)} = \frac{2 \log_2(3)}{\log_2(3)} = 2 \text{ pesées.}$$

3 - Algorithme en  $l_{\max} = 2$  pesées.  
et schéma 1.



## Résolution de Q2 :

$$H(Q_2) = \log_2(18) \quad 18 \text{ états car } 3 \text{ pièces & chaque pièce peut être soit + lourde, soit + légère.}$$

$$= \log_2(9 \times 2)$$

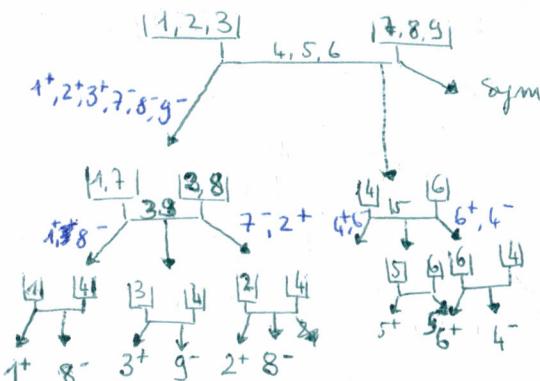
$$= 1 + \log_2(9) = 4,16 \text{ bits.} \quad H_{\text{équip}} = \left\{ \begin{array}{l} 1^+, 2^+, \dots, 9^+ \\ 1^-, 2^-, \dots, 9^- \end{array} \right\} \quad 18 \text{ états équip.}$$

2.  $\boxed{1,2,3} \quad \boxed{1,2,8,9}$   
 $\boxed{4,5,6}$

2. Question annexe ? 2 pesées sont-elles suffisantes pour résoudre Q2b = trouver 1 intrus parmi 6 et 6 pièces à 2 état  
 $\rightarrow$  sont état +/- lourd?

$$H(Q_{2b}) = \log_2(12) > 2 H_{\text{max pesée}} = 2 \log_2(3) = \log_2(9) \quad \text{NON mais pb mal posé.}$$

Car en fait 6 états possibles : 3 lourdes / 3 légères -



$$l_{\text{max}} = 3.$$

$$l_1 = l_2 = l_3 = \dots = l_{18} = 3.$$

Ds tout les cas

Balance numérique:  $H_{\text{pesée}} = \log_2(10^4)$  ( $1 \text{ kg par pas } 0,1 \Rightarrow 10^4 \text{ états}$ )

$$H(Q) = \log_2(5) \ll H_{\text{pesée}}$$

$\rightarrow$  on peut trouver en 1 pesée en mettant 1 pièce 1<sup>er</sup> sac, 1 pièce 2<sup>esac</sup>, ... 4 pièce, 5<sup>esac</sup>

## 3. Détection de la langue par comptage des lettres

① Détection langue 1 si

$$\text{Distance}(f_{\text{langue}}(n), P_c^{(1)}) < \text{Distance}(f_{\text{langue}}(n), P_c^{(2)})$$

Rap:  $\text{Distance}(x, y) = \sum_{i=1}^N |x_i - y_i|^2$  on veut minimiser la proba d'erreur.

1.ois. Vérifier formule  $P_c(z)$

4 alternatives possibles

	$\hat{H}_0$	$\hat{H}_1$
$H_0$	.	$\times$
$H_1$	$\times$	.

2. Réécriture de  $P_e$  à partir de  $\Pr(z|H_e) \stackrel{\text{notat}}{=} P_e(z)$  en f<sup>n</sup> seulement de  $P_e$ .

$$P(H_1|H_0) =$$

# TD : TI

3. Règle de décision optimale :  $P_e \min$

pour 1 mot  $z$  donné, règle de décision :

si  $C_1(z) < 0$  alors décision  $H_1$

si  $C_1(z) \geq 0$  alors décision  $H_2$

+ formellement :  $D_1 = \{z \in \Omega, C_1(z) < 0\}$

4.  $\Omega$  contient  $26^N$  élts.

Précalcul à l'avance de  $D_1$  et  $D_0$  serait très complexe, exp. avec  $N$ .

4 bis. Règle équivalente de décision.

$$D_1 = \left\{ z \in \Omega, \frac{P_1(z)}{P_0(z)} > \frac{P_{MO}}{P_{MI}} \right\}$$

5. suppose indépendance entre caractères successifs.

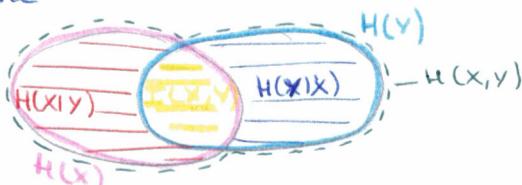
$$\begin{aligned} P_0(n) &= \prod_{k=1}^{26} P_0^{(0)} \\ &= \prod_{k=1}^{26} \left( P_c^{(0)} \right)^{m_k} \\ &\text{c'est } \uparrow \\ &\text{26 lettres} \end{aligned}$$

$$m_c = N \times f_c(m) \quad | \quad \begin{array}{l} \text{nb d'app. de la} \\ \text{lettre c ds le message m.} \end{array}$$

$$\begin{aligned} P_1(n) &= \prod_{k=1}^{26} P_1^{(1)} \\ &= \prod_{k=1}^{26} \left( P_c^{(1)} \right)^{m_k} \\ &\text{c'est } \uparrow \end{aligned}$$

## 4. Codage d'un couple de variables aléatoires

$X, Y$  : 2 va



$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \geq 0 \\ &= H(Y) - H(Y|X) \quad \text{ssi } X, Y \text{ indép} \\ &= H(X) + H(Y) - H(X, Y) \end{aligned}$$

$X_1, X_2 \in A_x = \{-1, +1\}$     $X_1, X_2$  indép.  
 $P_X = \beta_2, \beta_2$

A. Rvto de type 1 :  $\underline{X} = (X_1; X_2)$

$$\begin{aligned} a. H(X_1) &= \frac{1}{2} \log \frac{1}{2} + \frac{1}{2} \log \frac{1}{2} = \log_2(2) = 1 \text{ bit} \\ H(X_2) &= 1 \end{aligned}$$

$$b. H(X_2 | X_1) \stackrel{\text{indép}}{=} H(X_2) = 1 \text{ bit}$$

$$c. H(X) = H(X_1, X_2) \stackrel{\text{indép}}{=} H(X_1) + H(X_2) = 1+1 = 2 \text{ bits}$$

$$d. I(X;Y) \stackrel{\text{indép}}{=} 0$$

B. Rvto de type 2  $S = (S_+, S_-)$  avec  $S_+ = X_1 + X_2$  et  $S_- = X_1 - X_2$

Alphabet

$$S_+ : \{-2, 0, 2\} \quad S_- : \{-2, 0, 2\}$$

$$\begin{matrix} 1/4 & 1/2 & 1/4 \\ 1/4 & 1/2 & 1/4 \end{matrix}$$

$S_+$	$S_-$
0	-2
0	0
2	0
2	2

} m loi car symétrie  
du pb.

$$H(S_+) = \frac{1}{4}$$

b.  $H(S_+) = H(S_-) = -\frac{1}{2} \log \frac{1}{4} - \frac{1}{2} \log \frac{1}{2} = +\frac{1}{2} (\log 2^2 + \log 2) = -\frac{1}{2} (3 \log 2) = \frac{3}{2}$  bit d'information

c.  $H(X_1) < H(S^+) < H(X)$        $H(S^-) = 1,5$

2. Caract. de  $S = (S_+, S_-)$   
 $P_{S-|S+}$ :

<del>S<sup>-</sup></del>		-2	0	2	
Sachant S <sup>+</sup>		-2	0	1	0
		0	$\frac{1}{2}$	0	$\frac{1}{2}$
		2	0	1	$\frac{1}{2}$

$P(S_+, S_-) = P_{S-|S+} \cdot P(S^+)$

<del>S<sup>-</sup></del>		-2	0	2	
S <sup>+</sup>		-2	0	$\frac{1}{4}$	0
		0	$\frac{1}{4}$	0	$\frac{1}{4}$
		2	0	$\frac{1}{4}$	0

c.  $H(S) = H(S_+, S_-) = -\sum p(s^+, s^-) \log_2 p(s^+ | s^-)$   
 $= H(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$   
 $= \log_2(4) = 2$  bits

• entropie conditionnelle :  $H(S_-|S_+) = -\sum_{s^+ \text{ et } s^-} p(s^+, s^-) \log_2 p(s^- | s^+)$   
 $= -\frac{1}{4} \log_2(1) - \frac{1}{4} \log_2(\frac{1}{2}) - \frac{1}{4} \log_2(\frac{1}{2}) - \frac{1}{4} \log_2(1)$   
 $= 0,5$  bits

$H(S^-|S^+) = \sum p(s^+) H(s^-|S^+=s^+)$

$(S^-|S^+) = P(S=-2) H(S^-|S=-2) + P(S=0) H(S^-|S=0) + P(S=2) H(S^-|S=2)$

$I(S^+|S^-) = H(S^+) + H(S^-) - H(S^+|S^-) = 1,5 + 1,5 - 2 = 1$

e. Source X à N états  
 $Eff(S) = \frac{H(X)}{\log_2(N)}$  Redondance  
 $Red(S) = 1 - Eff(S)$   
 $Red(S) = 1 - \frac{H(X)}{\log_2(N)}$        $0 \leq Red(S) \leq 1$

$\leftarrow Red(S) = 1 - \frac{H(S)}{\log_2(133)} \approx 1 - \frac{2}{6}$

## 5.2. Codage de source binaire et ternaire.

### 1. Code binaire $D=2$

$$C = \{c_1, c_2, c_3, c_4, c_5\}$$

$$l_1=2 \quad l_2=2 \quad \dots \quad l_5=4$$

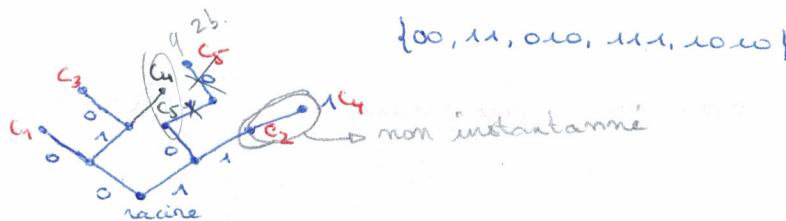
a) Déchiffrable? Non :  $111010 \rightarrow s_4, s_3 \rightarrow$   
ou  $s_2, s_5$ .

b) Instantané? Non car inst  $\Rightarrow$  dech. (non dech  $\Rightarrow$  non inst)

Rappel: Inst  $\Leftrightarrow$  Code qui satisfait la cond° du préfixe : aucun mot-code n'est le préfixe d'un autre mot code.

Ici :  $c_2=11$  est en préfixe de  $c_4=111$ .

Exemple de code instantané :



Question subsidiaire: Existe-t-il un code inst. avec un jeu de long? ( $l_1=l_2=2, l_3=l_4=3, l_5=4$ )

\* oui construit d'un code

\* Kraft Mac-Millan: CNS d'existence d'un code inst :  $\sum_{i=1}^N \left(\frac{1}{D}\right)^{l_i} \leq 1$

$$\text{ici : } 2 \times \left(\frac{1}{2}\right)^2 + 2 \cdot \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 \stackrel{13}{\leq} 1 \text{ ok.}$$

2.  $C = \{00, 11, 010, 011, 101\}$

a) instantané oui car CP ok (donc  $\Rightarrow$  déchiffrable)

b) si  $c_5 = 1010 \rightarrow c_5' = 10$  Code global inst & court VPs. cf m

Rq: Ici Kraft saturé :  $\sum_{i=1}^N \left(\frac{1}{D}\right)^{l_i} = 1$

3.  $P_S = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$  pour  $D=3$  code ternaire

a) avec  $l_i = l \ \forall i=1, \dots, S$  (long. fixe) on doit vérifier  $D^l \geq N : 3^l \geq 5$  ( $l_{\min}=2$ )

b)  $H(S) = - \sum_{i=1}^S p_i \log_2(p_i)$

$$= -\frac{1}{3} \log_2 \left(\frac{1}{3}\right) \times 2 - \frac{1}{3} \log_2 \left(\frac{1}{3}\right) \times 3 \quad H(S) = \frac{4}{3} \log_2(3)$$

$$= \frac{2}{3} \log_2 \left(\frac{3}{2}\right) + \frac{3}{3} \log_2(3) \times 2$$

\* red(S) =  $1 - \frac{H(S)}{\log_2(D)} = 1 - \frac{\frac{4}{3} \times 1,58}{2,32} \approx 9\%$

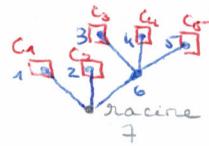
c)  $I_{\min} = \frac{H(S)}{\log_2(D)} = \frac{\frac{4}{3} \log_2(3)}{\log_2(3)} = \frac{4}{3} \approx 1,33$

taille de l'alphabet de codage

1<sup>er</sup> th. de Shannon: Il existe de codage tq le long moyen de code  $J = \sum p_i l_i$   
tq  $J = I_{\min} + \varepsilon \ (\forall \varepsilon > 0)$

Cours: Il existe de codage tq  $I_{\min} \leq J_k \leq I_{\min} + \frac{1}{k}$  lorsque codage par blocs de k lettres.

i) Mesures	no	P	NL	P	NL	P
$s_1$	1	1/3	1	1/3	7	1
$s_2$	2	1/3	2	1/3		
$s_3$	3	1/3	6	1/3		
$s_4$	4	1/9				
$s_5$	5	1/9				



$$A_k = \{0, 1, 2\}$$

$C_1 = 0$	1
$C_2 = 1$	1
$C_3 = 20$	2
$C_4 = 21$	2
$C_5 = 22$	2

ii)  $\overline{J} = \sum_i p_i l_i = (1 \times \frac{1}{3}) \times 2 + (2 \times \frac{1}{3}) \times 3 = \frac{4}{3}$

Efficacité du code :  $\text{Efficacité} = \frac{\overline{J}_{\text{min}}}{\overline{J}} = 1 = 100\%$

Code Huffman tjs optimal (si on peut pas trouver mieux) mais ici il est optimal absolu ( $\overline{J}=\overline{J}_{\text{min}}$ )

Conditions :  $P_i = \left(\frac{1}{d}\right)$  avec  $d$

Distribution D-adéquate.

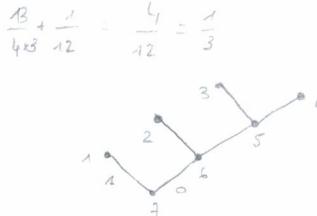
iii) Lorsque  $\overline{J} = \overline{J}_{\text{min}}$  : elts U après codage ( $\Rightarrow$  source U d'entropie maximale)

\*  $A_k = \{0, 1, 2\}$

et  $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$  loi uniforme ( $\Rightarrow U$  = source simple & de loi uniforme . suites d'elts indép.)

## 5.3. Longueurs des mots d'un code de Huffman

	P	NL
1	1/3	1/3 1/3 1/3 1/3 1/3 1/3 1/3
2	1/3	1/3 1/3 1/3 1/3 1/3 1/3 1/3
3	1/4	1/4 1/4 1/4 1/4
4	1/12	



$$A_k = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{12}$
---------------	---------------	---------------	----------------

$$H(S) = -\sum_{i=1}^k p_i \log_2 p_i \approx 1,855$$

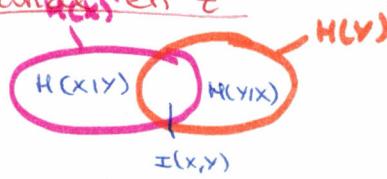
$$\overline{J}_{\text{min}} = \frac{H(S)}{\log_2(2)} \approx 1,855 \text{ digits/lettre}$$

Huffman :  $\frac{1}{3} \times 1 + 2 \times \frac{1}{3} + 3 \times \frac{1}{4} + 3 \times \frac{1}{12} = 2 \text{ digits/lettre}$

messages	1	2	3	4
$p_i$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{12}$
$l_i^* = \log_2(\frac{1}{p_i})$	1,58	1,58	2	3,58
li code Shannon	2	2	2	4
li Huff 1	1	2	3	3
li Huff 2	2	2	2	2

# TD : th. de l'info

## 7. Capacité de quelques canaux particuliers.



$$X \xrightarrow{(n)} Y$$

$$C = \max_{p_x} I(X, Y)$$

$$I(X, Y) = H(Y) - H(Y|X)$$

$P_X$	$X$	$Y$	$P_Y$
$P_1(X=0) = 1-p_1$	0	0	$P(Y=0)$
$P_1(X=1) = p_1$	1	0	$P(Y=1)$
		1	$P(Y=1   X=1)$

$\pi = \begin{bmatrix} 1 & 0 \\ 0,1 & 0,9 \end{bmatrix} \rightarrow \begin{array}{l} \text{ligne } 0 \\ \text{ligne } 1 \end{array}$

\*  $H(Y)$ ? en f'm de  $P_1$  ici  $N=2$ ,  $n=2$

\*  $H(Y|X)$ ?

$$* I(p_1) * \frac{dI(p_1)}{dp_1} = 0 \rightarrow C = I(p_1 = p_1^*) = I_{\max}$$

$$* H(Y) = -P(Y=0) \cdot \log_2(P(Y=0)) - (1-P(Y=0)) \log_2(1-P(Y=0))$$

$$\text{notab} H_2(P(Y=1)) \approx H_2(0,9 p_1)$$

$$* H(Y|X) = \sum_{i=0}^1 p(X=i) \cdot H(Y|X=i)$$

$$= P(X=0) \cdot H(Y|X=0) + P(X=1) \cdot H(Y|X=1)$$

$$= (1-p_1) \cdot H_2(1-p_1) + p_1 \cdot H_2(0,9 p_1)$$

$$= (1-p_1) H_2(1) + p_1 H_2(0,9)$$

$$H(Y|X) = H_2(0,1) \times p_1 = 0,469 p_1$$

$$\text{avec } H_2(0,1) \stackrel{\text{def}}{=} -0,1 \log_2(0,1) - 0,9 \log_2(0,9) \approx 0,469$$

$$\text{d'où } I(p_1) = H(Y) - H(Y|X) \quad I(p_1) = H_2(0,9 p_1) - 0,469 p_1$$

$$* p_1^* \text{ tq } \frac{dI(p_1)}{dp_1} = 0$$

$$\frac{dI(p_1)}{dp_1} = \frac{dH_2(0,9 p_1)}{dp_1} = 0,469$$

$$\frac{dH_2(p)}{dp} = \log_2\left(\frac{1-p}{p}\right) \quad \frac{dH_2(kp)}{dp} = k \cdot \log_2\left(\frac{1-kp}{kp}\right)$$

$$\frac{dI(p_1)}{dp_1} = 0,9 \log_2\left(\frac{1-0,9 p_1}{0,9 p_1}\right) - 0,469$$

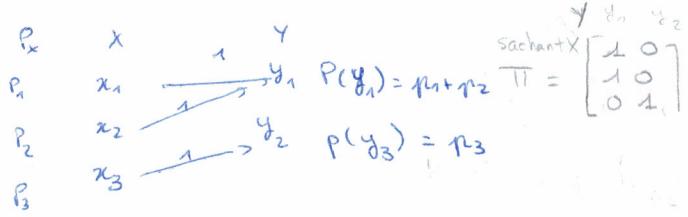
$$\frac{dI(p_1)}{dp_1} = 0 \Rightarrow \log_2\left(\frac{1-0,9 p_1}{0,9 p_1}\right) = \frac{0,469}{0,9} \approx 0,52$$

$$\Rightarrow \left(\frac{1-0,9 p_1}{0,9 p_1}\right) = 2^{0,52} \approx 1,434$$

$$p_1^* = 0,457 \Rightarrow p_0^* = 0,543$$

$$C = I(p_1 = 0,457) = 0,76 \text{ bit (par symbole)}$$

## Canal déterministe



1<sup>e</sup> méthode: En choisissant  $P_x = \left\{\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right\}$

$$\frac{1}{2}x_1 \longrightarrow y_1 \quad \text{UB sans bruit } I=1 \quad \text{or } C \leq \log_2(2) = 1 \text{ d'où } C = \max I = 1$$

$$\frac{1}{2}x_2 \longrightarrow y_2$$

car  $N=2=n$

$$H(Y) = H_2(P_3)$$

$$H(Y|X) = H(Y|X=i) \quad \forall \text{ ligne } i$$

$$= H_1 \text{ ligne } \text{cte indép de } P_x$$

$$\stackrel{\text{d'où}}{=} H_2(i) \stackrel{\text{d'où}}{=} 0$$

W/entrée

$$C = \max_{P_x} \{H(Y)\} - H_1 \text{ ligne}$$

$$C = \max_{P_3} H_2(P_3) - 0 = 1 \text{ avec } P_3^* = \frac{1}{2} \quad \forall P_1^* + P_2^* = \frac{1}{2}$$

## Canal sans équivoque

sachant  $Y$ :

$$\Pi = \begin{bmatrix} y_1 & y_2 & y_3 & y_4 \\ 0,5 & 0,5 & 0 & 0 \\ 0 & 0 & 0,5 & 0,5 \end{bmatrix}$$

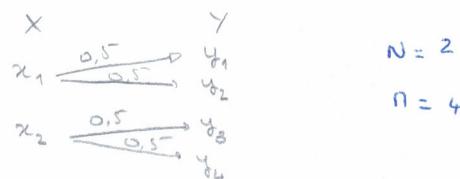
$$H(Y|X) = H(Y|X=i) \quad \forall \text{ ligne } i$$

$$= H_1 \text{ ligne } \text{cte indép de } P_x$$

$$= H(1/2, 1/2, 0, 0) = H_2(1/2) = 1$$

$$C = \max_{P_x} \{H(Y)\} - H_1 \text{ ligne}$$

$$= \max_{P_x} \{H(Y)\} - 1$$



$$\text{or si } P_1^* = \frac{1}{2} \Rightarrow H(Y) = 2 = \text{max absolu}$$

$$C = 2 - 1 = 1 \quad \text{avec } P_1^* = \frac{1}{2}$$

ou directement connu :

C 2<sup>ble</sup> uniforme

$$C = \log_2(n) - H_1 \text{ ligne avec } P_x \text{ uniforme}$$

# TD : TI

## 8- Principe du codage / décodage

### A- Canal continu gaussien.

Modèle d'observation.

$$y_k = (k + b_k) \sim \mathcal{N}(0, \sigma^2)$$

$$\in \{-1, 1\} \quad b_k \text{ indép densité de proba } p_b(a) = f_{0, \sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}$$

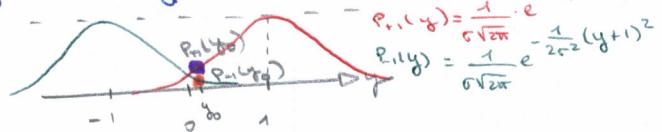


1. Modèle à  $k=0$   $y = c + b$

Vraisemblance de l'observation  $y$   $P_c(y)$  ou  $f(y|c)$  p.  $c=-1$  ou  $c=1$   $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-c)^2}{2\sigma^2}}$

Pour  $c$  fixé,  $y \sim \mathcal{N}(c, \sigma^2)$

$$P_c(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(y-c)^2}$$



2-  $P_e$  en fonction de  $Z_+$

Règle de décision : si  $y \in Z_+ \Rightarrow \hat{c} = +1$   
si  $y \notin Z_+ \Rightarrow \hat{c} = -1$

Test hyp binaire :

4 alternatives :  $(c = \pm 1) \cap (\hat{c} = \pm 1)$

probabilité qu'on ait choisi  $\hat{c}$  alors que  $c$  était  $\tilde{c}$ .

$$P_e = P_{\tilde{c}}(\hat{c}=1 | c=-1) \cdot p_{-1} + P_{\tilde{c}}(\hat{c}=-1 | c=1) \cdot p_{+1}$$

$$= P_{\tilde{c}}(y \in Z_+ | c=-1) \cdot p_{-1} + P_{\tilde{c}}(y \in \bar{Z}_+ | c=1) \cdot p_{+1}$$

$$= P_{\tilde{c}}(y \in Z_+ | c=-1) \cdot p_{-1} + 1 - P_{\tilde{c}}(y \in Z_+ | c=1) \cdot p_{+1}$$

$$= p_{+1} + \int_{y \in Z_+} (p_{-1}(y)p_{-1} - p_{+1}(y)p_{+1}) dy$$

3- Partition  $Z_+$  qui permet  $P_e$  min.

$$Z_+ = \{y \in \mathbb{R} \mid u(y) \leq 0\} \Leftrightarrow \text{décision } \hat{c} = +1 \text{ si } u(y) \leq 0 \\ \Leftrightarrow p_{-1}(y) \cdot p_{-1} - p_{+1}(y) \cdot p_{+1} \leq 0$$

$$\bullet \quad \frac{p_{+1}(y)}{p_{-1}(y)} \geq \frac{p_{-1}}{p_{+1}} \quad \text{- décision } \hat{c} = +1$$

$$\ln e^a = a$$

$$\ln e^{axb} = arb$$

Règle optimale :

$$Z_+ = \{y \in \mathbb{R} \mid LLR(y) \geq \log \frac{p_{-1}}{p_{+1}}\} \quad \text{LLR}(y) \stackrel{\text{def}}{=} \log \frac{p_{+1}(y)}{p_{-1}(y)}$$

3 bis - Trouver seuil optimal pour  $y$  ?

Exprimer d'abord LLR( $y$ ) en fonction de  $y$ .

$$LLR(y) = \ln \left( \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2} d^2(y, 1)}}{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2} d^2(y, -1)}} \right) = \frac{-1}{2\sigma^2} \left( d^2(y, 1) - d^2(y, -1) \right)$$

$$LLR(y) = \frac{-1}{2\sigma^2} (-4y) = \frac{2y}{\sigma^2}$$

Décision  $\hat{c} = +1$  si  $\frac{2y}{\sigma^2} \geq \ln \left( \frac{p_{-1}}{p_{+1}} \right)$

$$\text{ssi } y \geq \frac{\sigma^2}{2} \ln \frac{p_{-1}}{p_{+1}}$$

seuil opti sur  $y$

$$4 - P_+ = P_-$$

Décision  $\hat{c} = 1$  si  $y \geq 0$  et  $\hat{c} = \text{sgn}(y)$

$$\mathcal{Z}_+ = \{y \in \mathbb{R}, \text{LLR}(y) \geq 0\}$$

$$\text{LLR}(y) \geq 0 \Leftrightarrow \frac{1}{2\sigma^2} (\text{d}\epsilon^2(y_{i+1}) - \text{d}\epsilon^2(y_{i-1})) \geq 0$$

$$\Leftrightarrow \text{d}\epsilon^2(y_{i+1}) \leq \text{d}\epsilon^2(y_{i-1})$$

Avec symbole équiprobable :  $p^- = p^+ = \frac{1}{2}$

Règle max S;  $\rightarrow \hat{c} = \underset{c \in \{-1, 1\}}{\text{Arg Max}} P_c(y)$

$\hookrightarrow$  Règle min distance

$$\rightarrow \hat{c} = \underset{c \in \{-1, 1\}}{\text{Arg Min}} d_c^2(y, c)$$

6- Pb linéaire bits  $\rightarrow$  mot code  $\underline{c}^\circ$  = vecteur de n elt  $\in \{-1, 1\}$

bit 1  $\rightarrow$

$$\underline{c}^\circ$$

$$y = \underline{c} + \underline{B}$$

$$\text{d}\epsilon^2(y, c) = \sum_{i=0}^{m-1} (y_i - c_i)^2$$

$$B_+ \subset \mathbb{R} \rightarrow y \rightarrow \overline{(\text{sign})} - z \in \{-1, 1\}$$

$$1. \underset{c \in \{-1, 1\}}{\text{Arg Min}} \underline{c} + \underline{B}$$

$$\underline{\pi} = \begin{bmatrix} 1-p & p \\ p' & 1-p' \end{bmatrix}$$

$$p = P_c(\hat{c} = +1 | c = -1)$$

$$= P_c(y > 0 | c = -1)$$

$$= P_c(c + B > 0 | c = -1)$$

$$= P_c(-1 + B > 0) = P_c(B > \frac{1}{p})$$

$$= P_c(U > \frac{1}{p}) = Q(\frac{1}{p})$$

$$U \sim \mathcal{C}(0, 1) \text{ où } Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{u^2}{2}} du$$



De m<sup>e</sup>  $p' = P_c(\hat{c} = -1 | c = 1) = p$  d'au CBS.

2. Code par répétition :  $n = 2s + 1$  fois

bit 0  $\rightarrow c = -1 \rightarrow \underline{c}^\circ = -1, -1, \dots, -1$

bit 1  $\rightarrow c = +1 \rightarrow \underline{c}^1 = +1, +1, \dots, +1$

Modèle :

$$\underline{c} : \underbrace{\text{CBS}}_{\in \{\underline{c}^\circ, \underline{c}^1\}} \quad \underline{z} = \underbrace{\text{exp}}_{\sim \mathcal{N}(0, I)} \quad \underline{z} = 11-11-11-11-1 \dots$$

Vraisemblance ici ?

$$P_c(z | \underline{c}) = \overset{\text{note}}{P_c(z)} \text{ où } c \in \{\underline{c}^\circ, \underline{c}^1\}$$

$$P_c(z | \underline{c}) = (1-p)^{n-d_H(z, c)} p^{d_H(z, c)}$$

$$= (1-p)^n \left( \frac{p}{1-p} \right)^{d_H(z, c)}$$

$$d_H(z, c) = \text{nb d'elem diff. entre } z \text{ et } c$$

nb d'erreurs si  $\underline{c}$  est émis (?)

$$\hat{c} = \underline{c}^\circ \text{ si } d_H(\bar{z}, \underline{c}^\circ) < d_H(\bar{z}, \underline{c}^1) \Leftrightarrow \text{nb de } -1 < \text{nb de } 1$$

# TD 9 : Code de Hamming

Rappels: \* Codage par blocs :

$$\text{App: } \underline{\mathbf{u}}^k \xrightarrow{\text{F}_2} \underline{\mathbf{c}}^n \quad \begin{array}{l} n \geq k \\ \text{ou } \underline{\mathbf{F}}_2 \in \{0,1\}^{\text{dits}} \end{array}$$

$\underline{\mathbf{u}} = [u_1 u_2 \dots u_{k-1}]$        $\underline{\mathbf{c}} = [c_1 c_2 c_3 \dots c_n]$   
 $k$  digits       $n$  digits  
 message d'info      mot-code

1 code = 1 jeu de  $2^k$  mot-codes choisis parmi  $2^n$  possible.

\* Rendement du code  $R = \frac{k}{n}$       taux de codage  $1-R$   
 ou de redondance

\* Code linéaire:  $C(n, k)$

→ génération  $\underline{\mathbf{c}} = \underline{\mathbf{u}} \times \underline{\mathbf{G}}$       où  $\underline{\mathbf{G}} = \begin{bmatrix} \underline{\mathbf{g}}_1 \\ \vdots \\ \underline{\mathbf{g}}_k \end{bmatrix}$  matrice génératrice du code C.

$k$  vecteurs  $\underline{\mathbf{g}}_1, \dots, \underline{\mathbf{g}}_k$  forment une base génératrice du code.

On travaille avec + ou  $\oplus$

$$\begin{array}{c|cc} \oplus & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}$$

Si  $\underline{\mathbf{u}} = [u_1 \dots u_k]$  alors  $\underline{\mathbf{c}} = \underline{\mathbf{u}} \cdot \underline{\mathbf{G}} = u_1 \underline{\mathbf{g}}_1 + u_2 \underline{\mathbf{g}}_2 + \dots + u_k \underline{\mathbf{g}}_k$

→ Matrice de contrôle (de parité)

$$\underline{\mathbf{H}} = \begin{bmatrix} h_1 \\ \vdots \\ h_{n-k} \end{bmatrix}$$

Ptes:  $\underline{\mathbf{g}}_i \cdot \underline{\mathbf{h}}_j^T = 0 \quad \forall i \in \{1, k\} \quad \forall j \in \{1, n-k\}$

$$\underline{\mathbf{G}} \cdot \underline{\mathbf{H}}^T = 0$$

$$\underline{\mathbf{C}} \cdot \underline{\mathbf{H}}^T = [0 \dots 0]$$

→ Contrôle en réception à partir de mot reçu  $\underline{\mathbf{r}} (= \underline{\mathbf{c}} + \underline{\mathbf{e}})$  vect. erreur

Calcul syndrome:  $\underline{\mathbf{s}} = \underline{\mathbf{r}} \cdot \underline{\mathbf{H}}^T$

$$\& \underline{\mathbf{s}} = 0 \Rightarrow \underline{\mathbf{r}} \in \text{Code}$$

$\underline{\mathbf{s}} \neq 0 \Rightarrow \underline{\mathbf{r}} \notin \text{Code} \rightarrow$  présence d'erreurs.

Ex: Code de Hamming  $C(n, k) = \text{code linéaire à degré particulier } n = 2^m - 1$

Ex: Contrôle Global Info Rendement

$m$	$n = 2^m - 1$	$k = n - m$	$R = \frac{k}{n}$
2	3	1	1/3
3	7	4	4/7
4	15	11	11/15

$$1. R = \frac{k}{n} = 1 - \frac{m}{2^m - 1} \xrightarrow[m \rightarrow +\infty]{} 1$$

$$\lim_{m \rightarrow +\infty} R = 1$$

2) Code systématique:  $\forall \underline{c} \in \text{Code} \Rightarrow \text{mot code } \underline{u} = [\underline{c} | \underset{k=4}{\underbrace{\dots}} | \underset{n-k=3}{\underbrace{\dots}}]$

$$\underline{c} = \underline{u} \times \underline{G}$$

$$\text{avec } \underline{G} = \begin{bmatrix} \underline{I} & \underline{P} \end{bmatrix}_{4 \times 4} \quad 4 \times 3$$

comme

$$\underline{H} = \begin{bmatrix} \underline{P}^T & \underline{I} \end{bmatrix}_{3 \times 7} \Rightarrow \underline{G} \underline{H}^T = [\underline{I} | \underline{P}] = \underline{P} \oplus \underline{P} = \underline{0}$$

$$\text{ex: } \underline{g}_1 \cdot \underline{h}_1^T = [1 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0] \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}_{3 \times 7} = 2 + 1 = 3$$

$$\text{b) } \underline{H} = \left[ \begin{array}{ccc|cc} 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 \end{array} \right] \quad \underline{H}^T = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \underline{P} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

3. Message  $\underline{u} \rightarrow \text{Mot-code} \rightarrow \underline{c} = \underline{u} \cdot \underline{G} = (\text{dans } \underline{u} \underline{P})$

00000	00000	00000	$\rightarrow \underline{u} \underline{P}$
00001	00001	00001	$\rightarrow \text{Poids } = 1$
00010	00010	00010	$\rightarrow \text{Poids } = 2$
00011	00011	00011	$\rightarrow \text{Poids } = 3$
01000	01000	01000	$\rightarrow \text{Poids } = 4$
01001	01001	01001	
01100	01100	01100	
01110	01110	01110	
10000	10000	10000	
10001	10001	10001	
10100	10100	10100	
10111	10111	10111	
11000	11000	11000	
11001	11001	11001	
11100	11100	11100	
11111	11111	11111	

$$4. d_{\min} = \min_{\substack{i \neq j \\ \underline{c}_i, \underline{c}_j \in \text{code}}} d_H(\underline{c}_i; \underline{c}_j)$$

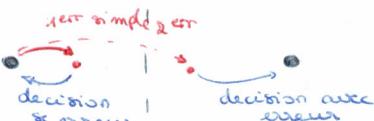
$$\text{avec } d_H(\underline{c}_i, \underline{c}_j) = \sum_{l=1}^7 c_i(l) c_j(l) = \text{Poids } (\underline{c}_i \oplus \underline{c}_j) = \text{nbr de } 1 \text{ ds le mot } \underline{z}$$

$$d_{\min} = \min_{\substack{i \neq j \\ \underline{c}_i, \underline{c}_j \in \text{code}}} \text{Poids } (\underline{c}_i \oplus \underline{c}_j)$$

$$d_{\min} = 3$$

$$\text{comme } \underline{z} = \min_{\substack{\underline{z} \in \text{code} \\ \underline{z} \neq 0}} \text{Poids } (\underline{z})$$

5. Capacité de correction:  $cc = \left\lfloor \frac{d_{\min}-1}{2} \right\rfloor = \left\lfloor \frac{3-1}{2} \right\rfloor = 1$  / 6- Ex: 1 message  $\underline{u} = 00100$   
 $\rightarrow$  capable de corriger les erreurs simples.



$$\text{Capacité de détection: } d_{\min} - 1 = 2$$

Canal:

Décodage:  $R =$

6. Ex: 1 message  $\underline{u} = 0010$

Tx mot-code  $\underline{c} = 00100011$

$$\begin{array}{l} \text{Canal : mot reçu } \underline{x} = \underline{c} + \underline{\varepsilon}_3 \\ \underline{x} = 000\ 0011 \end{array} \quad \left| \begin{array}{l} \text{erreur en 3<sup>e</sup> position} \\ \underline{\varepsilon}_3 = [00100000] \end{array} \right.$$

Décodage:  $\underline{x} = 000\ 0011$

Syndrome:  $\underline{s} = \underline{x} \cdot \underline{H}^T = 6^e + 7^e$  ligne de  $\underline{H}^T$

$\underline{s} = 011 \neq 0$  ) détecte (au moins) une erreur

décision: 3<sup>e</sup> col de  $\underline{H}^T$

en 3<sup>e</sup> lign de  $\underline{H}^T$

→ on décide de corriger la 3<sup>e</sup> pos.

$$\text{En effet, } \underline{s} = \underline{x} \cdot \underline{H}^T = (\underline{c} + \underline{\varepsilon}_3) \cdot \underline{H}^T = \underbrace{\underline{c} \cdot \underline{H}^T}_{\text{car } c \in \text{Code}} + \underline{\varepsilon}_3 \cdot \underline{H}^T$$

$$\underline{s} = \underline{\varepsilon}_3 \cdot \underline{H}^T = \text{i.e lign de } \underline{H}^T.$$

Code de Hamming: + code cc d'erreur p. corriger erreurs simples.

$$\therefore \dim m = n - k$$

2 cas distinguables  $\geq m$  pos. possibles + 1 cas pos.  
d'erreurs simples d'erreur

# TD W: Conséquence du théorème du codage canal

## 1. Révisions:

Source: Seq. de VA.

$$\underline{X} = X_1, X_2, \dots, X_\infty$$

(instant / indice)

$X_i$  in loi  $P_X$  ∀ instant  $i = 1, 2, \dots, \infty$

Source simple binaire

$$A_X = \{0; 1\}$$

$$P_X = \{0,185; 0,815\}$$

(D)

$X \rightarrow$  1 bit →  $V$  digits words

$$\text{Red}(X) = 1 - \frac{H(X)}{\log_2(\text{Card}(X))}$$

redondance

$$\text{Red}(X) = 0 \Leftrightarrow H(X) = \text{max possible}$$

### a) Les types de redondance d'une source:

- non indépendance entre éléments successifs ( $\Rightarrow$  source  $\neq$  simple = avec mémoire, dépendance)
- loi des  $X_i$  non uniforme

### b) ici source simple

$$H(X) = -0,865 \log_2(0,865) - 0,135 \log_2(0,135) \\ = 0,571 \text{ bit / lettre binaire de la source.}$$

$$\text{Red}(X) = 1 - \frac{H(X)}{\log_2(2)} = 1 - H(X) \approx 43\%$$

### c) $V_{\min}$ . Codage binaire (instantané)

1er th de Shannon: Il existe un procédé de codage permettant  $V = V_{\min} + \varepsilon \quad \forall \varepsilon > 0$  désiré

$$V_{\min} = \frac{H(X)}{\log_2(D)} \quad V_{\min} = \frac{H(X)}{\log_2(2)} = 0,571 \text{ digits par lettre.}$$

### d) Approche de $V_{\min}$ en pratique?

↪ codage par blocs de  $k$  lettres Huffman

$$\text{Garantie: } V_{\min} \leq V < V_{\min} + \frac{1}{k}$$

↪ ensemble de coder par blocs ( $k=1$ )

Si source a distribution D-adique:  $p(x_i) = \left(\frac{1}{D}\right)^{\text{entier li}} \Rightarrow li^* = -\log_2(p(x_i))$

Ici: Huffman sur:  $X$ :

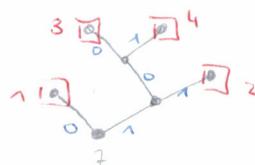
$$\begin{array}{ll} 0 \rightarrow 0 & V_{\min} = 1 \text{ digit / lettre} \\ 1 \rightarrow 1 \end{array}$$

### Huffman par extension d'ordre 2:

$$A_{X_2} = \{00, 01, 10, 11\}$$

$$P_{X_2} = 0,865^2; 0,865 \times 0,135; 0,135^2 = 0,748225; 0,116775; 0,116775; 0,018225$$

Résultat:



Codage:

00	→	0
01	→	11
10	→	100
11	→	101

$$V_1 = 1 \times 0,748225 + \dots + 3 \times 0,018225$$

$$V_2 = 1,3867 \text{ digits / mots de 2 lettres}$$

$$\tilde{V}_2 = \frac{V_2}{2} = 0,6933 \text{ digits / lettre.} < 1$$

$$\text{Eff. code} = \frac{V_{\min}}{\tilde{V}_2} = \frac{0,571}{0,693} \approx 82,3\%$$

$$V_{\min} \leq \tilde{V}_2 \leq V_{\min} + \frac{1}{2}$$

0,571      0,693      1,071

e) Code  $V \approx V_{\min}$  ( $\Rightarrow \text{Red}(U) \approx 0$ )  
 après codage  $H(U)$  mais absolue  
 $H(V)$

$\Leftrightarrow$  est après codage  
 \* successif ind.  
 \* de loi uniforme  
 $\Leftrightarrow U$  est source simple uniforme.

$$\begin{array}{c} X - \overline{|CS|} - U \\ \epsilon_{\{0,1\}} \end{array} \quad \xrightarrow{\text{CBS}} \quad \overline{|U|} \quad \epsilon_{\{0,1\}}$$

$$\Pi = \left[ \begin{array}{cc} 1-p & p \\ p & 1-p \end{array} \right] \quad \text{U/l'entrée} \quad 2^{\text{bit}} \text{ uniforme}$$

$\text{U/sortie}$

$$C = \log_2(n) - H_{\text{ligne}}$$

taille alphabet de sortie du canal.

CBS  $n=2$ :

$$(f_{\text{BS}} = 1 - H_2(p))$$

$$p = 0,08 \quad H_2(p) = 0,402 \quad \Rightarrow f_{\text{BS}} = 0,5978 \text{ bit/digit}$$

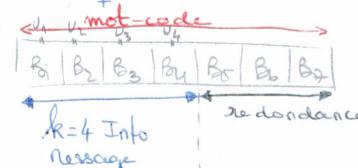
$$\text{Avec } H_2(p) = p \log_2(p) - (1-p) \log_2(1-p) \geq 0$$

2. Application du th. codage canal & code de Hamming

a) Avec code de Hamming  $C(\bar{n}; k)$

$$H(B) \text{ était } H(B) = \underline{k(B_1 \dots B_7)}$$

entropie moyenne par symbole



$$H(B) = \frac{k}{m} \cdot H(U) \quad \text{avec } H(U) \leq 1$$

$$H(B) = \frac{k}{m} = \frac{4}{7}$$

b) Rendement max selon le 2<sup>e</sup> th. de Shannon.

Un procédé codage/décodage Canal permettant l'appréciation de codage  $\leq \epsilon$ .

$$\forall \epsilon \geq 0 \text{ ssi } H(B) \leq C \Leftrightarrow \frac{k}{m} \leq C$$

c) Loi uniforme sur l'ensemble des mots codés?

$R_{\min} \Leftrightarrow$  Max Vrais ou Faux du (à faire du décodeur Valable si message équiprobable).

d) On suppose code Hamming  $(7, 4)$

\* Satisfaction Condition appliquée 2<sup>nd</sup> th. de Shannon  
 où Rendement  $\frac{4}{7} \approx 0,5714 < C = 0,5978$

\* On ne peut rien dire sur la loi pour mot  
 2<sup>e</sup> th. de Shannon = th. d'existence.

e) Parcours par mot après décodage avec Hamming  $(7, 4)$

Code de Hamming = Corriger erreurs simples. ( $d_{min} = 3 \Rightarrow c_c = 1$ )  
 $\forall n$

$P_{\text{err}} = \Pr(\text{Nb erreurs par mot} \geq 2)$

$$P_{\text{err}} = 1 - \Pr(\text{0 err/mot}) - \Pr(\text{1 err/mot})$$

$$= 1 - \underbrace{(1-p)^7}_{\text{Avec 6 erreurs}} - \underbrace{7(1-p)^6 \times p^1}_{\text{A cause où il y a une erreur}}$$

erreurs ss 0 ou 1

$p = 0,08$	$10^{-3}$	$10^{-7}$
$10,2\%$	$2 \cdot 10^{-5}$	$2 \cdot 10^{-13}$