Séance de TD01

Codage et compression

21 Mars 2022

(Codage de Hamming)

Exercice 01 : Après une transmission d'une information codée sur le canal selon l'algorithme de ***Hamming***, on a reçu ce message '**1110101**'. Vérifier si le message reçu est correcte ou non.

7	**6**	**5**	**4**	**3**	**2**	**1**
D_7	D_6	D_5	P_4	D_3	P_2	P_1
1	1	1	**0**	1	**0**	**1**

Étape 01 : vérifier les bits 1, 3, 5 et 7 ⇒ 1 + 1 + 1 + 1 = 4 (juste) ⇒ $A_1 = 0$.

Étape 02 : vérifier les bits 2, 3, 6 et 7 ⇒ 0 + 1 + 1 + 1 = 3 (faux) ⇒ $A_2 = 1$.

Étape 03 : vérifier les bits 4, 5, 6 et 7 ⇒ 0 + 1 + 1 + 1 = 3 (faux) ⇒ $A_3 = 1$.

Étape 04 : $A_3A_2A_1 = (110)_{\text{binaire}} = (6)_{\text{décimal}}$

⇒ L'erreur se trouve dans le bit numéro = 6 (on remplace '1' par '0').

Exercice 02 : vérifier cette information reçue '**1011100**'.

D7	D6	D5	P4	D3	P2	P1
1	**0**	**1**	**1**	**1**	**0**	**0**
7	6	5	4	3	2	1

Etape 1 : vérifier 1, 3, 5, 7 : P1 + D3 + D5 + D7 = 0+1+1+1=3 Donc c'est faux, A1=1.

Etape 2 : vérifier 2, 3, 6, 7 : P2 + D3 + D6 + D7 = 0+1+0+1=2 Donc c'est juste, A2=0.

Etape 3 : vérifier 4, 5, 6, 7 : P4 + D5 + D6 + D7 = 1+1+0+1=3 Donc c'est faux, A3=1.

Etape 4 : Lire $A_3A_2A_1$ en décimal pour localiser l'erreur et la corriger :

$A_3A_2A_1 = (101)_2 = 5$ (le bit erroné se trouve dans la position 5)

Donc, l'information correcte est '1001100'.

Exercice 03 : On veut coder ce message '**1110101**' selon l'algorithme de Hamming pour le transmettre sur un canal de transmission. Trouver le résultat.

Exercice 04 : Trouver le code de Hamming de ce message '**01011001**'

Séance de TD02

Codage et compression

15 Mai 2022

Exercice 01 :

Considérons une source discrète sans mémoire qui produit des symboles avec la loi de probabilité $p = (p0, p1, \cdots, p6) = \{0.4, 0.2, 0.15, 0.1, 0.05, 0.05, 0.05\}$ sur l'alphabet {a, b, c, d, e, f, g}.

1. Calculez H(p).

2. Trouvez le code de Huffman associé.

3. Calculez la longueur moyenne des mots code, et comparer-la avec H(p).

Exercice 02 :

Considérons une source discrète sans mémoire sur l'alphabet a0, a1, a2, a3, a4 qui produit selon la distribution de probabilité p0 = p2 = 0.2, p1 = 0.4, p3 = p4 = 0.1.

1. Calculez l'entropie de la source H(p).

2. Trouvez le code de Huffman associé.

3. Trouvez le code de Fano associé.

4. Calculez pour les deux codages les longueurs moyennes des mots, et comparez-les à H(p).

Exercice 03 :

Il existe plusieurs formats pour stocker un son. Par exemple :

A. Le format non compressé (WAV), où le son est enregistré en mono (un seul canal), l'amplitude du signal sonore codée sur 16 bits, échantillonnée à 44,1 kHz (1 Hz = 1 valeur par seconde).

B. Le format MP3, format WAV compressé pour pouvoir jouer en ne lisant que 128 kbits/s tout en conservant un signal sonore de bonne qualité. (Sachant que 1 kbits=1024 bits)

1. Quelle est la taille d'un fichier WAV standard mono de trois minutes en Mo ?

2. Quelle est la taille d'un fichier MP3 de trois minutes en Mo ?

3. Quel est le taux de compression du format MP3 ?

Exercice 04 :

Soit une source binaire X avec deux symboles '0' et '1' de probabilités respectives 'P' et '1-P'

1. Etudier l'état de son entropie H(X).
2. Déduire.

Exercice 05 :

Ce tableau montre les probabilités des symboles assimilées aux fréquences relatives d'apparition dans le message :

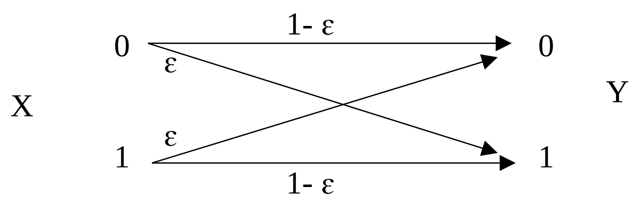
'espace'	A	B	E	G	I	L	S	T
1/10	1/10	1/10	1/10	1/10	1/10	2/10	1/10	1/10

Le codage arithmétique d'un message nous a donné : **0.2572167752**

Déduire le message original.

Exercice 06 :

Un canal est dit binaire symétrique ou (BSC) si p00 = p11 = 1-ε. (Un seul paramètre ε caractérise alors le canal).



1. Trouver la matrice stochastique de ce BSC sans erreurs.
2. Même question si l'incertitude est maximale.

Exercice 07 :

Un BSC reçoit des symboles à rs = 1000 sym/sec avec p00=p11=1/2. Calculer le Débit moyen (vitesse de transmission) pour ε=0.05, ε=0.3 et ε=0.5.

Exercice 08 :

Soit le code linéaire C(7,4) qui au vecteur d'information i associe le mot de code c tels que :

$i = (i_1\ i_2\ i_3\ i_4)$ et $C = (i_1\ i_2\ i_3\ i_4\ c_5\ c_6\ c_7)$ avec $c_5 = i_1 + i_3 + i_4$, $c_6 = i_1 + i_2 + i_3$ et $c_7 = i_2 + i_3 + i_4$.

1. Donner la matrice génératrice et la matrice de contrôle de ce code.
2. Soit i = (1 0 1 0), quel est le mot de code associé ?
3. Citer tous les mots de codes.
4. Calculer la distance minimale de ce code.
5. Combien d'erreurs peut-il détecter ?
6. Combien d'erreurs peut-il corriger ?
7. Soit le message reçu z = (1 1 1 1 0 0 1).
   a. D'après le récepteur, z est-il un mot du code C(7,4) (par démonstration) ?
   b. Effectuer une correction (si le mot z reçu est erroné) par la méthode du syndrome.

Correction TD02

Codage et compression

15 Mai 2022

Solution Exercice 01 :

1. H(p) =2.38
2. Code de Huffman :

Caractère	Code
a	0
b	100
c	101
d	110
e	1110
f	11110
g	11111

3. On obtient comme longueur moyenne du codage Huffman L = 2.45.

Solution Exercice 02 :

Solution Exercice 03 :

1. Taille du fichier WAV (en bit) = fréquence d'échantillonnage (en hertz) × quantification (en bit) × durée du fichier (en seconde) × nombre de voies (mono ou stéréo).

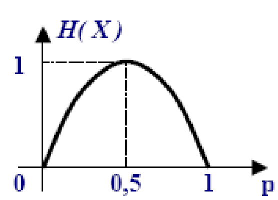
$= (44.1 \times 10^3) \times 16 \times (3 \times 60) \times 1$ (mono) = 121.12 Mbits = 15.14 Mo.

2. Taille du fichier MP3 (en bit) **=** 128 kbits × 3 × 60 (durée en second) = 22.5 Mbits = 2.85 Mo.

3. Taux de compression du format MP3 = $\frac{Taille\ originale - Taille\ compressée}{Taille\ originale} = \frac{15.14-2.85}{15.14} = 81.17\%$

Solution Exercice 04 :

1. H(X) = P.log(1/P) + (1-P).log(1/(1-P))
2. L'entropie H est maximale pour P = 0.5, et vaut zéro pour P = 0 ou P = 1. Aussi la variation de H est symétrique.



Solution Exercice 05 :

SYMBOLES	**Prob**	**Intervalles**
Espace (_)	1/10	[0 ,0.1 [
A	1/10	[0.1 ,0.2 [
B	1/10	[0.2 ,0.3 [
E	1/10	[0.3 ,0.4 [
G	1/10	[0.4 ,0.5 [
I	1/10	[0.5 ,0.6 [
L	2/10	[0.6 ,0.8 [
S	1/10	[0.8 ,0.9 [
T	1/10	[0.9 ,1]

Séquence	**Nouveau code**
B	0.2572167752
BI	0.572167752
BIL	0.72167752
BILL	0.6083876
BILL_	0.041938
BILL_G	0.41938
BILL_GA	0.1938
BILL_GAT	0.938
BILL_GATE	0.38
BILL_GATES	0.8
	0

Le message original = BILL GATES

Solution Exercice 06 :

On a un BSC :

1. Sans erreur = Sans bruit ⇒ P(X=i/Y=i) =1 ⇒ $P = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.
2. Incertitude maximale ⇒ P(X=i/Y=i) =1/2 ⇒ $P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$.

Solution Exercice 07 :

Débit moyen = *rs* * H(X) bits/sec

Pour ε = 0.05 ⇒ H(X) =

Solution Exercice 08 :

Soit le code linéaire C(7,4) qui au vecteur d'information i associe le mot de code c tels que :

i = (i_1 i_2 i_3 i_4) et C= (i_1 i_2 i_3 i_4 c_5 c_6 c_7) avec $c_5 = i_1 + i_3 + i_4$, $c_6 = i_1 + i_2 + i_3$ et $c_7 = i_2 + i_3 + i_4$.

1. La matrice génératrice G :

$$G = \begin{bmatrix} 1\,0\,0\,0 & 1\,1\,0 \\ 0\,1\,0\,0 & 0\,1\,1 \\ 0\,0\,1\,0 & 1\,1\,1 \\ 0\,0\,0\,1 & 1\,0\,1 \end{bmatrix}$$

La matrice de contrôle :

$$H = \begin{bmatrix} 1\,0\,1\,1\,1\,0\,0 \\ 1\,1\,1\,0\,0\,1\,0 \\ 0\,1\,1\,1\,0\,0\,1 \end{bmatrix}$$

2. Le mot de code associé à i = (1 0 1 0) ⇒ C = (1 0 1 0 0 0 1)

3. Les codes sont :

0000000 – 0100011 – 1000110 – 1100101 – 0010111 – 1010001 – 0110100 – 1110010 -

0001101 – 1001011 – 0101110 – 1101000 – 0011010 – 1011100 – 0111001 – 1111111.

4. La distance minimale *d* de ce code est *d* = 3

5. Nombre d'erreurs *ed* qu'il peut détecter : *ed* = d-1 ⇒ *ed* = 2

6. Nombre d'erreurs *ec* qu'il peut corriger : *ec* = (*d*-1)/2 ⇒ *ec* = 1

7. Soit le message reçu z = (1 1 1 1 0 0 1).

a. D'après le récepteur z n'est pas un mot du code c :

$z * H^T = (1111001) * \begin{bmatrix} 1\,0\,1\,1\,1\,0\,0 \\ 1\,1\,1\,0\,0\,1\,0 \\ 0\,1\,1\,1\,0\,0\,1 \end{bmatrix}^T = (110) \neq (000)$ ⇒ Donc z n'est pas un mot du code c.

b. Correction par la méthode du syndrome.

Erreur	**Syndrome**
1000000	110
0100000	011
0010000	111
0001000	101
0000100	100
0000010	010
0000001	001

$z * H^T = Er * H^T$

Donc Er = 1000000

Le mot reçu après correction est : $1111001 \oplus 1000000 = 0111001$