Cours Théorie de l'information

Tables de matières

Tables de matières

1. Information et codage

Les années 20 ont constitué le début du domaine de la théorie de l'information avec les travaux de Nyquist. Mais la théorie fut par les travaux du fameux ingénieur et mathématicien Shannon vers les fin des années 40.

1.1. But :

- ✓ Evaluer quantitativement le contenu d'un signal d'information
- ✓ déterminer la capacité d'un système de communication

1.2. Mesure de l'information

On appelle source un système capable de générer un flux d'information. La source sera continue ou discrète.
Nous allons nous focaliser sur milieu discret.

Soit X une source d'information dont l'alphabet est $\{x_1, x_2, \ldots\ldots x_m\}$
Si la source génère des symboles indépendants, alors elle est sans mémoire.

1.3. Quantité d'information

La quantité d'information représente une valeur d'information contenue dans chaque symbole qu'une source discrète est susceptible de générer.

$I(x_i) = -lb[Prob(x_i)]$ [Sh]

avec $Prob(x_i)$ la probabilité d'apparition de l'événement x_i.

1.4. Quantité d'information moyenne ou entropie

L'entropie est l'aptitude d'une source discrète à produire de l'information. Elle correspond à la moyenne des quantités d'informations de la source.

$H(X) = \sum_{i=1} Prob(x_i).I(x_i)$ [Sh/symbole]

1.5. La quantité de décision

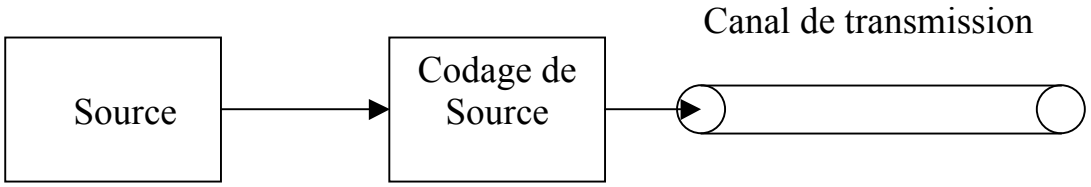
Elle correspond au maximum de l'entropie qui est atteint quand les symboles sont équiprobables.

D = lb(m) [bit/symbole]

1.6. Redondance

Ce paramètre exprime la différence entre la valeur de l'entropie et la quantité de décision

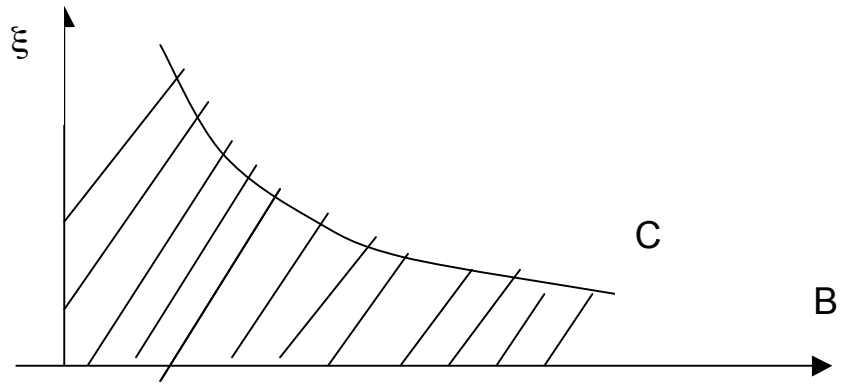
R = D – H



1.7. Capacité d'un canal :

Un canal de bande passante B en présence d'un bruit blanc gaussien a comme capacité C :

C=B.lb(1+ξ)



2. Compression et codage

2.1. But

- ✓ Diminuer la redondance
- ✓ Diminuer les erreurs

2.2. 1er Théorème de Shannon

Si H est l'entropie d'une source discrète sans mémoire, on peut théoriquement coder la source par une suite binaire en utilisant en moyenne H bits par symbole, sans jamais être inférieur à H.

Ce théorème a contribué à l'émergence de code réducteurs de redondance.

2.3. Codages sous optimaux

- o Codage de Fano-Shannon
- o Codage de Huffman
- o Codage arithmétique
- o Codage par blocs et plages

2.4. Code de Shannon-Fano

L'algorithme d'encodage suivant conduit à un code à décodage unique et instantané :

1) Ordonner les caractères à encoder selon l'ordre décroissant de leurs probabilité.
2) Diviser l'ensemble des caractères à encoder en deux sous-ensembles aussi équiprobables que possible
3) Attribuer à chaque sous-ensemble en symbole binaire distinct.
4) répéter la procédure pour chaque caractère à encoder, jusqu'à ce que chacun d'eux possède une transcription binaire distincte.

Exemple : n=nombre des caractères
n=8

X_i	Prob(x_i)	code					l_i	Prob(Xi).l_i
X_1	0,40	0	0				2	0,8
X_2	0,15	0	1				2	0,3
X_3	0,15	1	0	0			3	0,45
X_4	0,10	1	0	1			3	0,3
X_5	0,10	1	1	0			3	0,3
X_6	0,06	1	1	1	0		4	0,24
X_7	0,02	1	1	1	1	0	5	0,1
X_8	0,02	1	1	1	1	1	5	0,1

L = 2,59

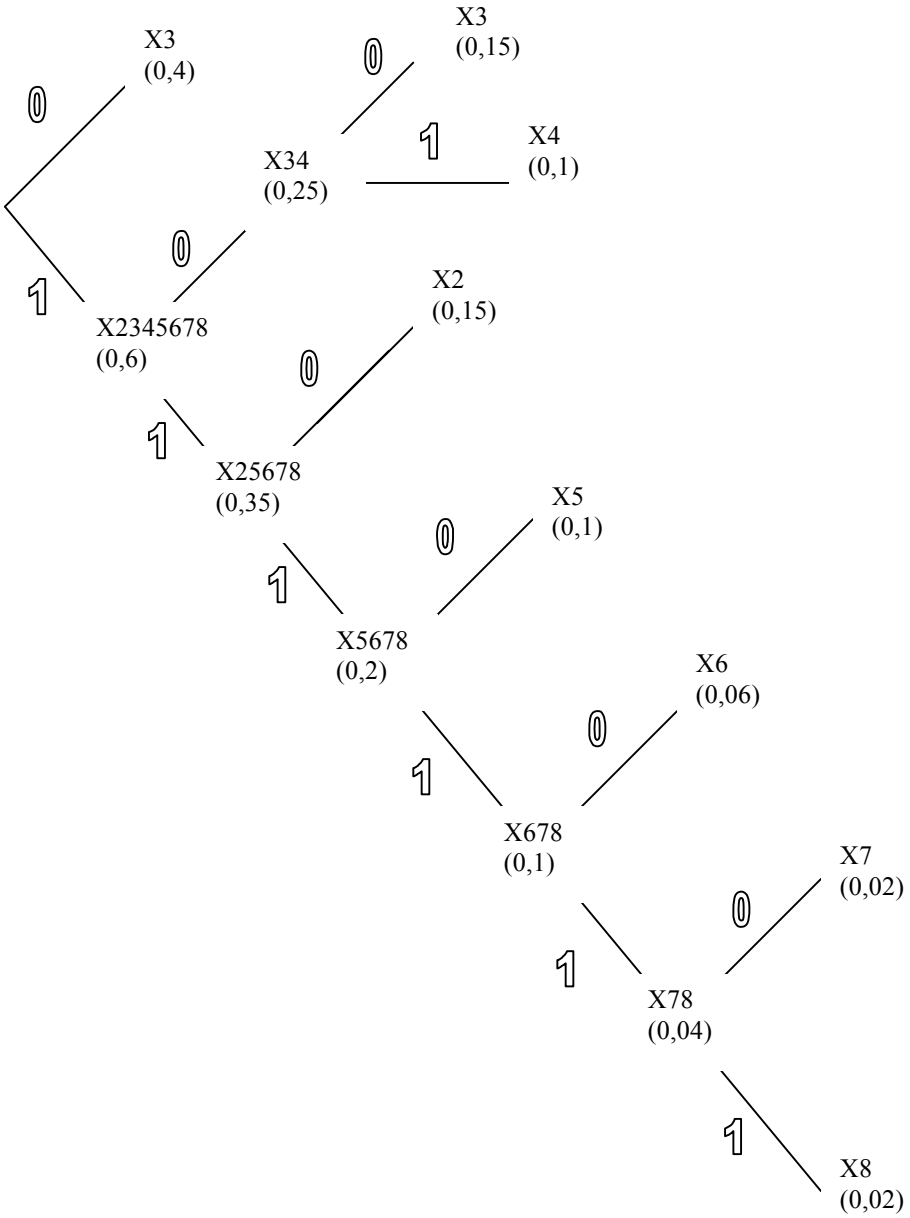
Entropie de la source : H = 2,485 Sh/symbole
Quantité de décision : D = 3 bits/symbole
Longueur moyenne du symbole : 2,59 bits/symbole
Redondance de la source : Rs = 0,515 bit/symbole
Redondance résiduelle : Rr = 0,105 bit/symbole

2.5. Code de Huffman

Cette algorithme est utilisé dans plusieurs applications (Fax, Codage H.261, etc.)

1) Ordonner les caractères à encoder selon l'ordre décroissant de leurs probabilité.
2) Construire un arbre d'encodage en partant de la fin. Ceci correspond à attribuer aux deux caractères à encoder de probabilité la plus faible (Xn et Xn-1) les symboles 0 et 1.
3) Remplacer dans la liste ces deux caractères par un caractère conjoint (Xn,n-1) de probabilité cumulée Prob(Xn-1) + Prob(Xn). Le nouveau caractère remplace un nœud de l'arbre.
4) Répéter la procédure en combinant à nouveau les deux caractères de plus faible probabilité.
5) Poursuivre ainsi jusqu'à ce que la probabilité totale soit égale à un. Le code se lit en lisant en arrière depuis le sommet final.

Exemple d'encodage selon l'algorithme de Huffman : soit à encoder la même source X que dans l'exemple précédent, illustrant la méthode de Shannon-Fano.



Arbre d'encodage de Huffman

X_i	Prob(x_i)		code					l_i	Prob(Xi) .l_i
X_1	0,40	0						1	0.40
X_2	0,15	0	1	0				3	0,45
X_3	0,15	1	0	0				3	0,45
X_4	0,10	1	0	1				3	0,3
X_5	0,10	1	1	1	0			4	0,4
X_6	0,06	1	1	1	1	0		5	0,3
X_7	0,02	1	1	1	1	1	0	6	0,12
X_8	0,02	1	1	1	1	1	1	6	0,12

L = 2,54

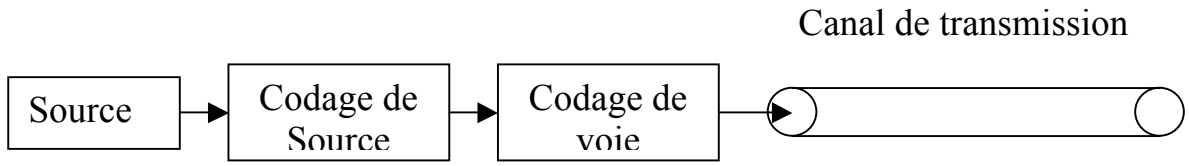
Entropie de la source : H = 2,485 Sh/symbole

Quantité de décision : D = 3 bits/symbole
Longueur moyenne du symbole : 2,54 bits/symbole
Redondance de la source : Rs = 0,515 bit/symbole
Redondance résiduelle : Rr = 0,1055 bit/symbole

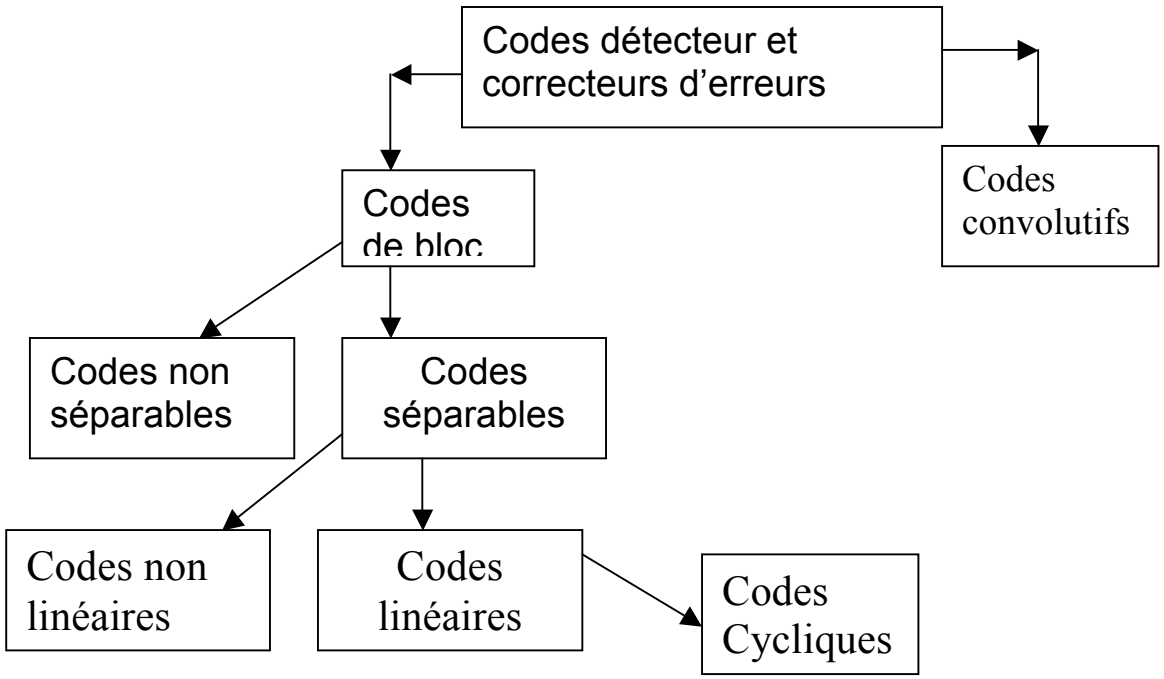
2.6.2ème théorème de Shannon

Si le débit d'une source est inférieur à la capacité C d'un canal en présence d'un bruit, on peut diminuer le taux d'erreurs autant que l'on veut moyennant codage.

L'existence de ce théorème a favorisé le développement des codages détecteurs et correcteurs d'erreurs. Le principe est simple il suffit d'introduire de la redondance pour pouvoir corriger ou détecter les erreurs.



2.7.Classification de code



3. Classification des codes

Il est intéressant d'utiliser un exemple pour expliquer la notion de classement d'un code. Soit le tableau suivant

xi	code 1	code 2	Code 3	Code 4	Code 5	Code 6
x1	00	01	0	0	0	1
x2	10	10	1	10	01	01
x3	11	11	00	110	011	001
x4	11	00	11	111	0111	0001

a. Code de longueur fixe

Les codes 1 et 2 sont de longueur fixe (2)

b. Code de longueur variable

Les codes 3, 4, 5 et 6 sont de longueur variables

c. Code univoque

Chaque mot code représente un seul symbole. C'est le cas par exemple des code 2, 4 et 5

d. Code sans préfixe

Un code sans préfixe est code dont chaque symbole n'est pas le préfixe d'un d'autre. C'est le cas des codes 2, 4 et 6

e. Code déchiffrable de façon unique

C'est un code qui permet de décoder les mots codes sans ambiguïté et de manière univoque. Le code 3 n'est pas déchiffrable de manière unique car la séquence reçue 000 peut être interprétée comme x1,x1,x1, x1,x3 ou x3,x1.

Une condition suffisante pour possède cette propriété est que le code soit sans préfixe ; exemple du code 5 qui est un code avec préfixe et il est déchiffrable d'une manière unique.

f. Code instantané

Un code à déchiffrage unique est dit code instantané si tout mot code est identifiable sans examen des symboles du mot de code qui suit. Les codes instantanés sont des codes sans préfixe.

g. Code optimal

Un code est optimal s'il est instantané et présente une longueur moyenne L minimale pour une source donnée.

3.1. Inégalité de Kraft

Soit X = {x_i, i=1….n} une source discrète amnésique et soit n_i la longueur du mot code de chaque symbole x_i.
Une condition nécessaire et suffisante d'existence d'un code binaire instantané a pour expression :

$$K = \sum_{i=1}^{m} 2^{-n_i} \leq 1$$

relation connue sous le nom de l'inégalité de Kraft.

4. Codage d'un canal de transmission

C'est l'étude de code qui permet d'acheminer avec fiabilité des informations dans un canal bruité. Cette étude est basée sur le théorème de Shannon qui affirme l'existence d'un codage sans donner les moyens de le construire.

4.1. Notions élémentaires de probabilités

Probabilité d'avoir pile ou face

Le calcul de la probabilité de changement d'un bit lors de la transmission revient à calculer la probabilité d'avoir pile ou face lors du jet d'une pièce.

La probabilité d'avoir deux fois pile si l'on lance une pièce de monnaie 3 fois de suites :

p(2xpiles/3jets) = p(PPF)+
p(PFP)+
p(FPP)

p(PPF)= p(PFP)=p(FPP) =p(config)=p(P).p(P).P(F)=$p(P)^2$.p(F)
Le nombre de configuration est donné par le nombre de combinaison de 3 éléments prix deux à deux. On a donc :
p(2xpiles/3jets)=C^3_2 . $p(P)^2$.p(F)

Exemples
On lance un dé 5 fois.

a. Quelle est la probabilité d'avoir 5 fois le chiffre 6 ?
b. Quelle est la probabilité de ne pas avoir le chiffre 6 ?
c. Quelle est la probabilité d'avoir une fois au plus le chiffre 6?

a. le dé à six faces donc 6 nombres x={1, 2, 3, 4, 5, 6}. la probabilité d'avoir le six si le dé n'est pas pipé c'est-à-dire il y a équiprobabilité est 1/6. p(5fois) = $(1/6)^5$
b. pn1=$(5/6)^5$
c. p= 1-$(1-1/6)^5$

Le binôme de Newton est donnée par

$$(x+y)^n = C_0^n.x^n + C_1^n.x^{n-1}.y^1 + C_2^n.x^{n-2}.y^2 + ... + C_n^n y^n = \sum_{i=0}^{n} C_i^n.x^{n-i}.y^i \; avec C_i^n = \frac{n!}{i!.(n-i)!}$$

Dans le cas d'une probabilité et son complément
$((1-p)+p)^5$ est égale à 1

4.2. Probabilité d'erreurs

Sur un canal bruité les bits risquent de se faire modifier indépendamment les uns des autres avec une probabilité p_b.
La probabilité de modifier m bits dans une séquence de N bits est :
$p(m \text{ erreurs}) = C^N_m . p_b^m .(1- p_b)^{N-m}$

$p(>=m \text{ erreurs}) = {}^N\sum_{k=m} C^N_k . p_b^k .(1- p_b)^{N-k}$

Dans la majorité des cas la probabilité d'une erreur sur un bit est faible $p_b<<1$.

$(1- p_b)^{N-k} \approx 1$ et $p_b^k >> p_b^{k+1}$
donc $p(>=m \text{ erreurs}) \approx C^N_m . p_b^m$

pour m =1 on a $p(>=1 \text{ erreur}) \approx C^N_1 . p_b^m = N.p_b$

4.3. Détection d'erreurs par bit de parité

Bit de parité simple

La manière la plus simple de détecter une erreur simple est d'utiliser le code de parité par l'ajout d'un bit de parité.
A la réception du message, on peut vérifier s'il y a eu lieu d'erreur y compris sur le bit de parité. Ce système détecte seulement les erreurs impaires car si deux bits changent en même temps la parité est conservée.
La probabilité d'erreurs non détectées est donc donnée par :

$$p_{en} = \sum_{i=2,pair}^{N} C_i^N . p_b^{\,i} .(1-p_b)^{N-i}$$

Si $p_b <<1$ alors $p_{en} = C_2^N . p_b^{\,2} = N.(N-1)p_b/2$

Un bit de parité améliore donc beaucoup la probabilité de détection si les erreurs ne sont pas corrélées entre elle.