MA387
Théorie de l'information et codage
TD

Nicolas Barbot

7 avril 2015

1 Théorie de l'information

1.1 Exercice

On tire une main de 4 cartes dans un jeu de 32 cartes. On définit les événements suivants :
- E_1 la main ne contient aucune carte en dessous du valet
- E_2 la main ne contient pas de figure
- E_3 la main contient 4 cartes de même nom
- E_4 la main contient les 4 as

Calculer $h(E_i)$, $i(E_1; E_3)$ et $i(E_1; E_2)$
Calculer la quantité d'information pour spécifier 4 cartes

1.2 Exercice

On lance une pièce biaisée de probabilité p. Calculer l'entropie de la source.
On lance un dé équilibré. Calculer l'entropie de la source.
Calculer l'information mutuelle entre la face du haut et la face du bas du dé.
Calculer l'information mutuelle entre la face du haut et celle face au joueur.

1.3 Exercice

Deux capteurs sont caractérisés par les probabilités de fausse alarme et de non détection suivantes :

	FA	ND
Capteur 1	0.05	0.03
Capteur 2	0.001	0.06

Les deux capteurs suivent la même source uniforme.
Quel est le meilleur capteur au sens de la théorie de l'information.
Refaire l'exercice avec une probabilité d'alarme de 1/100.

1.4 Exercice

On lance une pièce biaisée jusqu'a l'obtention d'une face. Soit X le nombre de lancers. Soit Y le nombre de lancers (total) jusqu'a l'obtention d'une seconde face.
Calculer $H(X)$.
Déterminer une borne sur $H(Y)$.

1.5 Exercice

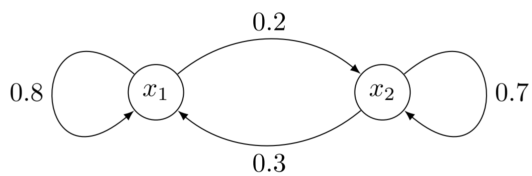
Soit X et Y deux variables aléatoire binaires indépendantes de probabilités respectives p_X et p_Y. Soit $Z = X \oplus Y$ leur somme modulo 2.
Déterminer p_Z, à quelle condition $p_Z = 1/2$.
Comparer $H(Z)$ avec $H(X)$ dans le cas ou $p_Y = p_X$.

1.6 Exercice

Alice lance un dé deux fois. Quel est le nombre minimale de question (oui/non) que Bob doit poser pour retrouver le nombre d'Alice.
Déterminer le nombre moyen minimal de questions.

1.7 Exercice

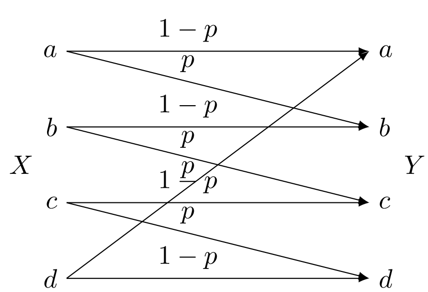
On considère la source binaire X suivante :



Déterminer l'expression de $H(X)$.
On donne $p(x_1) = 0.6$ et $p(x_2) = 0.4$ Calculer $H(X)$ et comparer avec une source sans mémoire de même probabilité.

1.8 Exercice

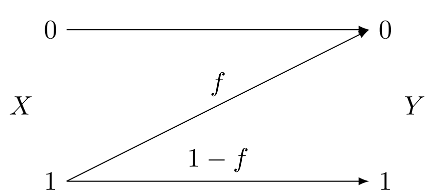
On considère le canal suivant :



Déterminer la capacité de ce canal ainsi que la distribution d'entrée optimale.

1.9 Exercice

On considère le canal suivant :



Déterminer la capacité de ce canal ainsi que la distribution d'entrée optimale.

1.10 Exercice

On considère le canal AWGN à entrée et sortie continues. Sachant que pour une puissance donnée, la densité de probabilité optimale (maximisant l'information mutuelle est gaussienne, calculer la capacité du canal AWGN.

1.11 Exercice

Un canal insère une erreur sur chaque bloc de 8 bits. L'erreur peut se situer sur chaque bit avec une probabilité uniforme.
Déterminer la capacité du canal.

2 Codage de source

2.1 Exercice

Une source possède un alphabet de 4 lettres. On souhaite encoder ces symboles en utilisant des mots de codes de longueurs 1, 2, 2 et 3 bits. Déterminer s'il existe un code préfixe permettant de satisfaire cette contrainte.

2.2 Exercice

Un code est défini par le mapping suivant :

a_i	$c(a_i)$	p_i	l_i	$h(p_i)$
a	00	0.3		
b	01	0.3		
c	10	0.2		
d	110	0.15		
e	111	0.05		

Déterminer si le code est un code préfixe.
Compléter le tableau.
Pour une source caractérisée par la distribution p_i, déterminer la longueur du code.
Pour la même source, calculer l'information pour chaque symbole ainsi que l'entropie de la source.

2.3 Exercice

En utilisant le code de l'exercice précédent, encodé la séquence :
abacbbaacdaabababbcabbe
Décoder la séquence :
0011010111101000101000101011000011010100110.
La séquence issue d'un codeur par symbole est dépourvue de redondance mais est extrêmement sensible aux erreurs de transmission. Décoder la même séquence en ajoutant une erreur sur le 3eme bit.

2.4 Exercice

Une source possède un alphabet de 8 lettres $(a_1, \ldots, a_8)$ de probabilités respectives (0.26, 0.20, 0.15, 0.12, 0.10,0.08 0.05,0.04).
Utiliser l'algorithme de Shannon Fano pour déterminer un code binaire pour cette source.
Utiliser l'algorithme de Huffman pour déterminer un code binaire pour cette source.
Déterminer le nombre moyen de bit par symbole dans chaque cas
Comparer ces valeurs à l'entropie de la source.

2.5 Exercice

On considère la source définie à l'exercice 1.7.
Déterminer le code de Huffman pour des symboles constitués de 2 lettres.
Déterminer le code de Huffman pour des symboles constitués de 3 lettres.
Calculer la longueur du code dans chaque cas.

2.6 Exercice

On considère la séquence :
01101001010101010100100100101000000100101010010000000100010101
Encoder cette séquence en utilisant l'algorithme de Lempel Ziv.

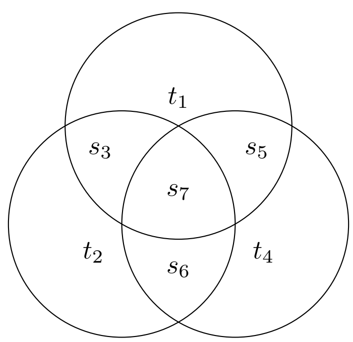
3 Codage de canal

3.1 Exercice

On considère le code R_3.
Déterminer le rendement du code R_3.
Détailler l'algorithme de décodage.
Déterminer la valeur de la probabilité d'erreur bloc p_B pour un canal binaire symétrique de probabilité f.
En déduire la valeur de la probabilité d'erreur bit p_b.
Même questions pour le code R_9.
En remarquant que l'encodage peut être réalisé en concaténant deux codes R_3, proposer un algorithme de décodage alternatif et estimer la probabilité d'erreur bit p_b dans les deux cas.

3.2 Exercice

On considère le code de Hamming (7,4).



Déterminer le rendement du code de Hamming (7,4).
Énumérer l'ensemble des mots de code et en déduire la distance minimale du code.
Déduire de la représentation graphique, l'expression de la matrice génératrice G et de la matrice de contrôle H du code.
Déduire à partir de G et H l'expression de la matrice génératrice systématique G_s et de la matrice de parité systématique H_s.
Sachant que le syndrome peut être calculé par $s = eH^T$ où e est le motif d'erreur, calculer tous les syndromes correspondant à un motif d'erreur de poids 1. En déduire une méthode permettant de décoder un message reçu.
Décoder les blocs suivants :

$$r = 0100101001011001110111111110$$

Calculer la probabilité de non détection de ce code.
En remarquant que le code de Hamming ne peut corriger que des erreurs simples, déterminez la probabilité d'erreur bloc p_B exacte sur le CBS de probabilité f ainsi qu'une valeur approchée (approximation au premier ordre). En déduire une expression du taux d'erreur binaire p_b.

3.3 Exercice

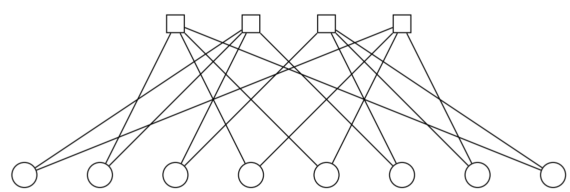
On considère le code convolutif (3,5,7)
Dessiner la structure de l'encodeur.
Représenter le code sous forme de chaîne de Markov.
Encoder la séquence 1001101011101011101.
Mêmes questions avec le code (15,17).
La mission Galileo vers Jupiter a utilisé le code (46321,51271,63667,70535).
Déterminer le nombre d'états.

3.4 Exercice

On considère le code convolutif (7,5,7).
Encoder la séquence 101100.
Lors de la transmission sur un canal binaire symétrique, des erreurs se produisent sur les bits 6 10 13 et 14.
Décoder la séquence reçue avec l'algorithme de Viterbi.
Déterminer la distance libre du code.

3.5 Exercice

On considère le code LDPC suivant :



Déduire le rendement du code, ainsi que sa matrice de parité H.
Parmi les mots suivants, déterminer lesquels sont des mots de code :
10000000, 10100101, 10101010, 00011000, 00000000, 11111111.
Pour les mots qui ne sont pas des mots de code, appliquer l'algorithme de décodage bit flipping.
Dans ce graphe de Tanner, identifier un cycle de longueur 4.