Enoncé des exercices sur le Codage de Source

Exercice 1

Soit p la probabilité d'un événement, tracer la valeur de la quantité d'information relative à l'événement en fonction de p pour $0 \leq p \leq 1$.

Exercice 2

Une source émet aléatoirement un symbole parmi quatre symboles possibles. Ces quatre symboles ont des probabilités d'occurrence telles que $p_0 = 0.4$, $p_1 = 0.3$, $p_2 = 0.2$ et $p_3 = 0.1$ et sont statistiquement indépendants.

1. Calculer l'information associée à l'émission de chacun de ces 4 symboles.
2. Calculer l'entropie de la source.

Exercice 3

Considérons une source sans mémoire (les symboles émis sont statistiquement indépendants) dont l'alphabet est constitué par K symboles équiprobables.

1. Quelle est le meilleur codage possible pour une telle source, à longueur fixe ou variable ? Pourquoi ?
2. Quelle condition doit satisfaire K pour que l'efficacité du codage soit maximale ?

Exercice 4

Considérez les 4 codes listés dans la table suivante :

Symbole	Code I	Code II	Code III	Code IV
s_0	0	0	0	00
s_1	10	01	01	01
s_2	110	001	011	10
s_3	1110	0010	110	110
s_4	1111	0011	111	111

1. Lesquels de ces codes sont à décodage instantanés ?
2. Calculer l'inégalité de Kraft-McMillan pour chacun de ces codes. Discutez les résultats en fonctions de ceux obtenus en 1).

Exercice 5

Considérez des lettres d'un alphabet ayant les probabilités d'apparition telles que :

Lettre	a	i	l	m	n	o	p	y
Probabilité	0.1	0.1	0.2	0.1	0.1	0.2	0.1	0.1

TABLE 1.1 – Table associée à l'exercice 5

Calculez deux codes différents de Huffman pour cet alphabet. Dans un cas, reportez un symbole combiné dans l'arbre à la position la plus haute, dans l'autre dans la position la plus basse. Pour chacun des codes obtenus, calculez la longueur moyenne des mots code ainsi que la variance de la longueur moyenne des mots code. Lequel de ces deux codages possibles choisirez-vous en pratique ?

Exercice 6

Une source discrète sans mémoire a un alphabet de 7 symboles dont les probabilités d'apparition sont décrites dans la table suivante :

Symbole	s_0	s_1	s_2	s_3	s_4	s_5	s_6
Probabilité	1/4	1/4	1/8	1/8	1/8	1/16	1/16

Calculer le code de Huffman associé ainsi que son rendement.

Exercice 7

Considérez une source discrète sans mémoire avec un alphabet $\{s_0,\ s_1, s_2\}$ dont les probabilités d'apparition sont respectivement $\{0.7,\ 0.15{,}0.15\}$. Appliquez le codage d'Huffman à cette source et montrez que la longueur moyenne du code est $1.3\ bits/symbole$.

Exercice 8

En considérant la figure 1.1, donnez les codes associés aux symboles A, B, C, D, E, F, G.

Exercice 9

Un calculateur exécute 4 instructions qui sont représentées par les mots code (00, 01, 10, 11). En supposant que ces instructions sont utilisées de manière indépendante avec des probabilités (1/2, 1/8, 1/8, 1/4), calculez le pourcentage d'économie de bits qui pourrait être réalisé avec un codage de source optimal. Trouver un code de Huffman qui réalise ce codage optimal.

Exercice 10

Considérez la séquence binaire suivante

$$011100011100011100011100$$

Utilisez l'algorithme de Lempel-Ziv pour encoder cette séquence.

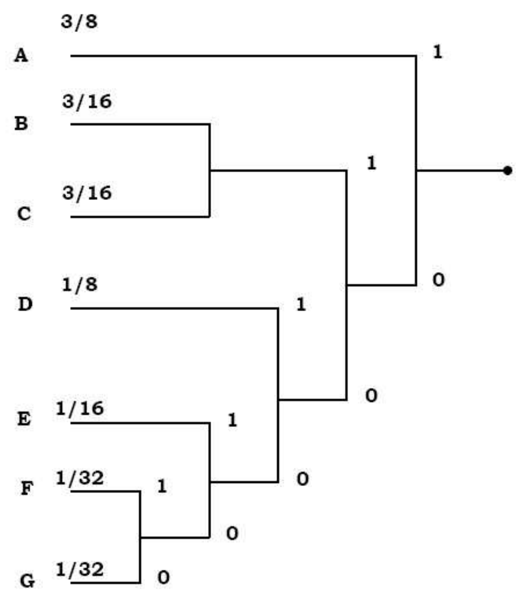

FIGURE 1.1 – Figure liée à l'exercice 8

Exercice 11

Une source sans mémoire a un alphabet $\mathcal{A}$ tel que

$$\mathcal{A} = \{-5,\ -3,\ -1,\ 0,\ 1,\ 3,\ 5\}$$

Les probabilités correspondantes sont

$$\{0.05,\ 0.1,\ 0.1,\ 0.15,\ 0.05,\ 0.25,\ 0.3\}$$

1. Trouver l'entropie de la source.
2. En supposant que la source est quantifiée selon la règle suivante :

$$\begin{array}{lllllll} Q(-5) & = & Q(-3) & = & -4 & & \\ Q(-1) & = & Q(0) & = & Q(1) & = & 0 \\ Q(3) & = & Q(5) & = & 4 & & \end{array}$$

   trouver l'entropie de la source quantifiée.
3. Proposer un codage optimal de cette source quantifiée.

Exercice 12

Une source sans mémoire émet des symboles $a_1,\ a_2,\ a_3,\ a_4$ avec des probabilités correspondantes $p_1 = p_2 = 0.3365$ et $p_3 = p_4 = 0.1635$.

1. Trouver l'entropie de la source.
2. Trouver un code de Huffman pour cette source.
3. Calculer la longueur moyenne du code obtenu.
4. Calculer le rendement du code obtenu.

Exercice 13

Une source sans mémoire a un alphabet de 5 symboles S_1, S_2, S_3, S_4 et S_5. Ces cinq symboles sont équiprobables. Evaluez le rendement d'un code binaire de longueur constante dans les trois cas suivants :

1. On code chaque symbole.
2. On code des paires de symboles.
3. On code des successions de trois symboles.

Exercice 14

Une source sans mémoire émet des symboles S_1, S_2, S_3, S_4 avec des probabilités $p_1 = 0.5$, $p_2 = 0.25$, $p_3 = 0.125$ et $p_4 = 0.125$.

1. Coder chaque symbole avec un code de longueur fixe et calculer le rendement du code.
2. Trouver un code à décodage instantané de rendement 100%.

Exercice 15

Le schéma de la Figure 1 montre un schéma d'automate de Markov. Cet automate produit un symbole toutes les T $s.$, ce symbole peut être s_0 ou s_1. On lit le schéma de la Figure 1.2 de la manière suivante ; si s_0 a été produit à l'instant $k.T$, la probabilité d'obtenir un nouveau s_0 à l'instant $(k+1).T$ est $1/4$, alors que la probabilité d'obtenir s_1 à $(k+1).T$ est alors $3/4$. (De tels automates sont utilisés pour modéliser la production de parole.). On observera qu'une tel automate produit des données qui sont dépendantes dans le temps. En supposant que l'on code des séquences de deux symboles consécutifs :

1. Calculer l'entropie de la source constituée par l'automate.
2. Proposer un codage de Huffman des séquences de deux symboles.
3. Calculer l'efficacité du code ainsi créé.

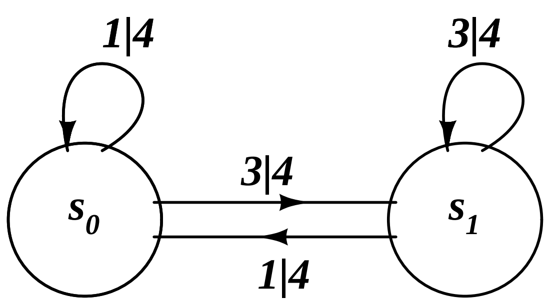

FIGURE 1.2 – Automate de Markov

Exercice 16

On considère une source S_1 sans mémoire délivrant des symboles $\{s_0,\ s_1,\ s_2,\ s_3\}$ avec des probabilités respectives $1/4,\ 1/2,\ 1/16,\ 3/16$. D'autre part on considère une source S_2 sans mémoire délivrant des symboles $\{t_0,\ t_1,\ t_2,\ t_3\}$ avec des probabilités respectives $1/8,\ 1/32,\ 1/64,\ 53/64$. On considère ensuite une source S_3 qui délivre aléatoirement soit un symbole de S_1, soit un symbole de S_2 à chaque coup d'horloge ; la probabilité que S_1 soit choisie est de $3/4$.

1. Calculer l'entropie de chacune des trois sources S_1, S_2 et S_3.
2. Proposer un codage de Huffman des séquences de symboles pour chacune des trois sources.
3. Calculer l'efficacité des trois codes ainsi créés.

Exercice 17

Soit une source S_1 sans mémoire délivrant des symboles $\{s_0,\ s_1,\ s_2,\ s_3,\ s_4\}$ avec des probabilités respectives $1/16,\ 1/16,\ 1/8,\ 1/4,\ 1/2$. Sans faire appel à un graphe de type Huffman, proposer un code à décodage instantané qui ait une efficacité de 100%. Justifier vos choix et expliquez bien pourquoi l'efficacité est de 100%.

Exercice 18

On considère la séquence suivante que l'on veut encoder par l'algorithme de Lempel-Ziv.

$$0101101110111101111101111111$$

On considère que 0 est le premier mot code et que 1 est le second mot code. On considère d'autre part que les mots code sont numérotés à partir de 1.

1. Calculer la séquence codée par l'algorithme de Lempel-Ziv.
2. Calculer le taux de compression obtenu.
3. Vérifiez que la séquence obtenue est bien juste en procédant au décodage de la séquence.