EISTI. **Année scolaire 2008-2009**

DÉPARTEMENT " INFORMATIQUE "

THÉORIE DE L'INFORMATION

Série d'exercices N^{o}4. CORRIGE.

PARTIE I. CODAGE DE HUFFMAN

Nous allons considérer une source d'alphabet

$$\Omega = \{a, b, c, d, e, f, g, h, i, j\}$$

et un canal d'alphabet binaire $\{0, 1\}$. Voici les définitions importantes.

Théorème de la borne inférieure de longueur de code

Théorème 0.1. *Soit une source S d'alphabet $\Omega_S = \{s_1, \ldots, s_n\}$ de taille n et de distribution de probabilités $P_S = \{p_1, \ldots, p_n\}$. Soit un canal d'alphabet binaire $\Omega_C = \{0, 1\}$ sans bruit, stationnaire et sans mémoire. Soit un code déchiffrable $\{m_1, \ldots, m_n\}$ de longueurs de mots $\{l_1, \ldots, l_n\}$.*
Alors la longueur moyenne de mots de code vérifie :

$$\bar{L} = \sum_{i=1}^{n} p(s_i) l_i \geq H(S)$$

L'égalité n'est possible que si $\forall i = 1, \ldots, n, \;\; p_i = 2^{-l_i}$.

Code absolument optimal C'est un code dont la longueur moyenne de mots est égale à la borne inférieure, $H(S)$.

Code optimal. Un code est dit optimal dans une certaine classe de codes si sa longueur moyenne de mots est minimale dans cette classe. La classe de codes la plus importante est celle de codes sans préfixe.

Exercice 1. Soit une source d'alphabet Ω et de distribution de probabilité suivante

s_i	a	b	c	d	e	f	g	h	i	j
m_i	0.2	0.05	0.1	0.05	0.15	0.05	0.1	0.05	0.15	0.1

1. Quelle est la longueur moyenne minimale pour un code binaire de cette source ?

$$H(S) = 3.14$$

2. Existe-t-il un code absolument optimal ? Non car les probabilités ne sont pas des puissances de $\frac{1}{2}$.
3. Construire un code sans préfixe optimal selon la méthode de Huffman.

s_i	a	e	i	c	g	j	b	d	f	h
m_i	0.2	0.15	0.15	0.1	0.1	0.1	0.05	0.05	0.05	0.05

s_i	a	e	i	c	g	j	fh	b	d	
m_i	0.2	0.15	0.15	0.1	0.1	0.1	0.1	0.05	0.05	

s_i	a	e	i	c	g	j	fh	bd		
m_i	0.2	0.15	0.15	0.1	0.1	0.1	0.1	0.1		

s_i	a	fhbd	e	i	c	g	j			
m_i	0.2	0.2	0.15	0.15	0.1	0.1	0.1			

s_i	a	fhbd	gj	e	i	c				
m_i	0.2	0.2	0.2	0.15	0.15	0.1				

s_i	ic	a	fhbd	gj	e					
m_i	0.25	0.2	0.2	0.2	0.15					

s_i	gje	ic	a	fhbd						
m_i	0.35	0.25	0.2	0.2						

s_i	afhbd	gje	ic							
m_i	0.4	0.35	0.25							

s_i	gjeic	afhbd								
m_i	0.6	0.4								

4. Représenter ce code sous forme d'arbre.

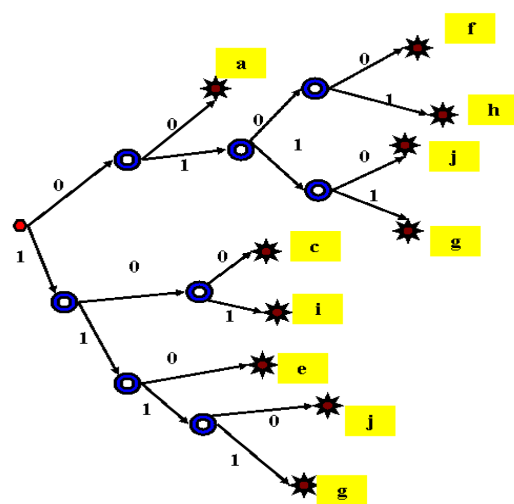

FIGURE 1 – Un code absolument optimal sans préfixe