

Techniques de Compression et Utilisation d'Outils Cryptographiques Web et Mobile

– Compression de données en informatique

TRAVAUX DIRIGES N° 1

TD Compression de données en informatique

Exercice 1 :

Soit une source qui émet 4 symboles x_1, x_2, x_3, x_4 et 4 codes C_1, C_2, C_3, C_4 tels que :

x	$p(x)$	C_1	C_2	C_3	C_4
x_1	0.5	0	0	0	0
x_2	0.25	0	1	10	01
x_3	0.125	1	00	110	011
x_4	0.125	10	11	111	0111

1. Rappeler la définition d'un code à décodage unique, d'un code préfixe. Quelle est l'avantage d'avoir un code à préfixe ?
2. Parmi ces 4 codes lesquels satisfont la condition de préfixe ? Pourquoi ?
3. Parmi ces 4 codes lesquels sont à décodage unique ? Pourquoi ?
4. Quel est le plus efficace ?

Exercice 2 :

Soit une source qui génère des lettres de l'alphabet $A = \{a_1, a_2, a_3, a_4, a_5\}$ avec les probabilités suivantes : $p(a_1) = 0.15$; $p(a_2) = 0.04$; $p(a_3) = 0.26$; $p(a_4) = 0.05$; $p(a_5) = 0.5$.

1. Calculer l'entropie de la source
2. Construire le code de Fano pour cette source. Calculer l'efficacité du code.
3. Construire le code de Huffman pour cette source. Calculer l'efficacité du code.
4. Comparer les deux codes ?
5. Dans chaque cas n'y-a-t-il qu'une seule manière de construire le code ?

Exercice 3 :

On considère le fichier $\mathcal{F} = \{BBDBFDFDDBBDEDCAFABB\}$.

On supposera que la distribution de probabilité des symboles est donnée par la fréquence relative d'apparition du symbole dans le fichier \mathcal{F} .

1. Calculez les probabilités des symboles A, B, C, D, E et F .
2. Calculez l'entropie H du fichier \mathcal{F} .
3. Déterminez l'encodage de Fano des symboles.
4. Déterminez l'encodage d'Huffman des symboles.
5. Quelle serait la taille minimale d'un fichier \mathcal{F}' , composé de 1000 symboles tirés du même dictionnaire que \mathcal{F} et avec la même distribution de probabilité ?
6. Quelle taille permet d'atteindre l'encodage de Fano pour \mathcal{F}' ?
7. Quelle taille permet d'atteindre l'encodage de Huffman pour \mathcal{F}' ?

Techniques de Compression et Utilisation d'Outils Cryptographiques Web et Mobile

– Compression de données en informatique

Exercice 4 :

On considère l'image ci-dessous, où les valeurs des pixels sont marquées. On considère que chaque pixel est codé sur 1 octet.

10	10	10	10	10
10	90	90	90	90
180	180	180	180	180
100	100	100	100	10
15	15	15	15	15

1. Quelle est la taille de l'image en nombre de bits.
2. Donner le codage de RLE et codage de Huffman.
3. Calculer la taille de l'image après chaque codage.
4. Quel est le taux de compression de chaque codage ?
5. En déduire le gain compression de chaque codage.

Exercice 5 :

1. Soit le message suivant : "ABCACABC" à encoder en LZW avec un dictionnaire initial réduit aux 5 premières lettres : dico[0] = 'A', dico[1] = 'B', dico[2] = 'C', dico[3] = 'D', dico[4] = 'E'.

a. Coder le message initial en remplissant le tableau ci-dessous :

étape	préc. au début du TQ	caractère lu : c	séquence	dans dico ?	décision	dico	code	préc. final
-------	----------------------	------------------	----------	-------------	----------	------	------	-------------

b. Décompresser cette séquence binaire en complétant le tableau ci-dessous.

étape	précédente	taille du dico ⇒ nbits	code	courante après test	c	nouvelle et dico
1	vide	5 entrées ⇒ code 3 bits	0/3	A	A	vide+A, dico[0] = A