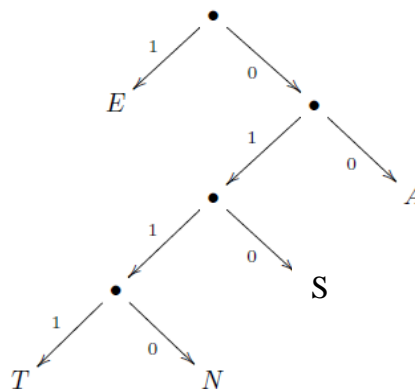


Correction TD 1 Partie2

Codage source

Exercice 1 : Codage source : Huffman

Soit l'alphabet {E; A;S;N ;T}. Le codage source suivant, donné comme arbre de Huffman :



1. Attribuer à chaque caractère son code binaire selon l'arbre de Huffman.

Réponse :

Il faut lire l'arbre du haut vers le bas

Alphabet	E	A	S	N	T
Code binaire	1	00	010	0110	0111

2.

- a. En utilisant l'arbre, déchiffrer le message $m=00011001111011001101010$.

Réponse :

Le message : 00011001111011001101010

⇒ ANTENNES

- b. Quelle sera la taille du codage du message en code ASCII (1 octet par caractère)?

Réponse :

Le message est de 8 caractères, chaque caractère est codé par 1 octet (ASCII) => La taille du codage du message sera donc 8 octets.

Taille du codage du message en ASCII= $8 * 8 = 64$ bits

c. Quelle sera la taille minimale du codage du message en code à longueur fixe ?

Réponse :

Le message (ANTENNES) est composé de 5 symboles.

Soit M le nombre de symboles et n le nombre de bit nécessaire pour coder un symbole:

$$n = \lceil \log_2(M) \rceil$$

$$2^2 \leq M=5 < 2^3 \Rightarrow n=3$$

Ou $n = \lceil \log_2(M) \rceil = \lceil \log_2(5) \rceil = \ln(5) / \ln(2) = 2.32 \Rightarrow n = 3 \text{ bits}$

Taille du message en code à longueur fixe = $8 \times 3 = 24$ bits.

d. Quel est le taux de compression Huffman ?

Réponse :

Taux de compression ?

Taille du message codé en ASCII est 64 bits

Taille du message codé en Huffman est 23 bits (00011001111011001101010)

$$\text{Taux de Compression Huffman} = \tau_1 = \frac{\text{Taille du message en Huffman}}{\text{Taille du message en ASCII}}$$

$$\Rightarrow \text{A.N } \tau_1 = \frac{23}{64} = 0.36 = 36\%$$

e. Quel est le taux de compression du codage à longueur fixe?

Réponse :

Taux de compression ?

Taille du message codé en ASCII est 64 bits

Taille du message codé en code à longueur fixe est 24 bits

Taux de Compression en code à longueur fixe

$$\tau_2 = \frac{\text{Taille du message en code à longueur fixe}}{\text{Taille du message en ASCII}}$$

$$\Rightarrow \text{A.N } \tau_2 = \frac{24}{64} = 0.37 = 37\%$$

f. Comparer et commenter les deux taux trouvés

Réponse :

$$\tau_1 < \tau_2$$

Le code de Huffman permet d'avoir une meilleure compression que le code à longueur fixe pour le message « ANTENNES »

3. a. Peut-on améliorer le taux de compression du code Huffman pour le message m ?

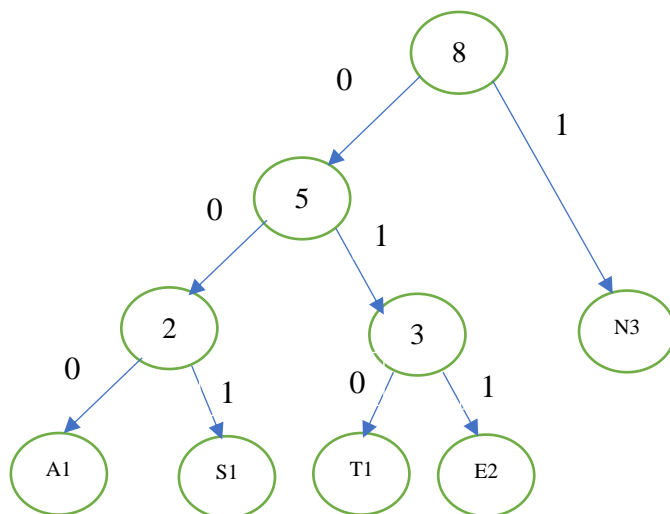
Réponse :

L'arbre de Huffman utilisé n'est pas optimal pour coder le message ANTENNES. En effet, le symbole 'N' a plus d'occurrence que les autres symboles. Il apparaît 3 fois sur le message. Alors que le code binaire proposé du symboles 'N' est sur plus de bits que les autres symboles.

b. Proposer une solution et recalculer le taux de compression dans ce cas.

Réponse :

N.B. : La solution proposée n'est pas unique. D'autres configurations d'arbre sont possibles !



Symbole	A	S	T	E	N
Nombre d'occurrence	1	1	1	2	3
Code binaire	000	001	010	011	1

Taille du message = 18 bits

$$\Rightarrow \tau_3 = \frac{18}{64} = 0.28 = 28\%$$

Exercice 2 (Codage source : Huffman)

On considère le codage de Huffman suivant :

$a \rightarrow 1$
 $b \rightarrow 011$
 $c \rightarrow 000$
 $r \rightarrow 010$
 $t \rightarrow 001$

1. Décodez le message $m = 011101000110111000$.

Réponse :

$m = 011101000110111000$

Le message est : **bartabac**

2. On suppose que le texte initial était codé en ASCII (1 octet par caractère). Quel est le gain en codant ce même texte avec le code de Huffman ci-dessus ?

Réponse :

La taille du message **m** est 8 caractères

Avec le codage ASCII la taille du message est : $8 \times 8 \text{ bits} = 64 \text{ bits} = 8 \text{ octets}$.

Alors qu'avec le codage Huffman **m** (011101000110111000) est de taille : 18 bits

$$\Rightarrow \text{Le gain du codage Huffman : } G = 1 - \tau = 1 - \frac{\text{Taille du message en Huffman}}{\text{Taille du message en ASCII}}$$

$$\underline{\text{AN:}} \quad G = 1 - \frac{18}{64} = 0.72$$

3. Lorsqu'on transmet un texte codé par cette méthode, on doit également transmettre le dictionnaire de décryptage.

a. Expliquer ce que cela peut signifier.

Réponse :

À la réception, le récepteur doit disposer le même dictionnaire de Huffman (défini et utilisé par l'émetteur) pour décoder son message correctement.

b. Quelles conséquences sur le taux de compression ?

Réponse :

L'expression du taux de compression devient :

$$\tau = \frac{\text{Taille du message en Huffman} + \text{taille du dictionnaire}}{\text{Taille du message en ASCII}}$$

Taux dans ce cas est plus élevé. Pour des fichiers de faible taille, il y a un risque que l'en-tête contenant le dictionnaire plus les données compressées dépassent la taille du fichier original.