

Examen TI

Exercice 1 : Entropie d'une source et codage binaire (5 pts)

Soit une variable aléatoire $X = \{a, b, c, d\}$

- Déterminer l'entropie $H(X_p)$ quand les probabilités associées à X sont $P(X) = \{\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}\}$
- Déterminer l'entropie $H(X_q)$ quand les probabilités associées à X sont $Q(X) = \{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\}$
- On définit le code binaire suivant :

$$c(X = a) = 0$$

$$c(X = b) = 10$$

$$c(X = c) = 110$$

$$c(X = d) = 111$$

Ce code est-t-il un code sans préfixe ?

Calculer la longueur moyenne des mots de code (en bits/symbole) quand X est distribuée selon P(X) puis selon Q(X).

- Comparer les résultats. Dans quel cas de distribution de probabilité ce code est absolument optimal ?

Exercice 2 : Modélisation mathématique d'un canal (5 pts)

Soit une source binaire $X = \{0, 1\}$ de probabilité $P(x=0)=0.5$ et $P(x=1)=0.5$.

Soit un récepteur (canal) de même alphabet, pour chaque symbole émis, le canal de transmission à la probabilité $\frac{1}{4}$ de faire une erreur.

- Représenter le canal par un graphe et en déduire la matrice de distribution des probabilités de réception $P(Y|X)$
- Trouver les distributions conjointe $P(X,Y)$ et marginale $P(Y)$

Nous rappelons la formule : $p(y|x) = \frac{p(x,y)}{p(x)} \Rightarrow p(x,y) = p(x) * p(y|x)$

- Calculer les entropies $H(X)$, $H(Y)$, $H(X,Y)$ et $H(Y|X)$
- Calculer l'information mutuelle de X et Y : $I(X,Y)$ et $I(Y,X)$

Exercice 3 : Compression RLE et Huffman (10 pts)

Considérant l'alphabet de la source $\mathbf{X} = \{a, b, c, d\}$ et la séquence de texte à compresser : Seq= «**abbcccdddd**».

Méthode 1 – Codage de Huffman

- Constituer la table des fréquences des symboles de l'alphabet X contenus dans la séquence donnée.
- Calculer l'Entropie associée. Que dire de l'efficacité du codage qui associe 8 bits à chaque caractère ?
- A partir de la table des fréquences, déterminer un arbre de Huffman et le code associé.
- Calculer la taille du code de Huffman (en bits) de cette séquence. Quel est le taux de compression?

Méthode 2 – Codage RLE (Run Length Encoding)

- Constituer le code RLE correspondant à la séquence Seq
- Chaque couple obtenue est encodée sur 16 bits : 8 bits pour le nombre de répétitions et 8 bits pour le caractère répété. Quelle est la longueur du code (en bits) obtenu ?
- Calculer le taux de compression.

Méthode 3 – Codage RLE + Huffman

- Appliquer le codage de Huffman au code RLE trouvé et calculer le nouveau taux de compression
- Comparer les performances des 3 méthodes (Huffman, RLE, RLE+Huffman).