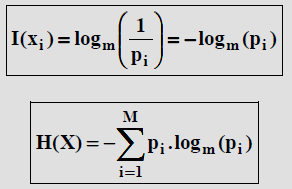
**Théorie de l’information et codage**

**TP N°1: Notion d’entropie et compression de l’information par codage de Huffman**

Les étudiants sont conduits à manipuler les notions relatives à la mesure de l’information et à l’application de la méthode de Huffman pour le codage et la compression de données.

1. **Mesure de l’information : Rappel**



On considère un alphabet de M symboles que nous notons : xi, avec 1≤ i ≤M. Chaque symbole xi apparait avec une probabilité pi=P(xi).

Le contenu informatif de chaque symbole xi peut alors être mesurée par une grandeur de la forme de I(xi). L’entropie associée à cet alphabet de M symboles est représentée par H(X)

1. Voici le texte suivant :

DIDONS DINA DIT ON DU DOS D’UN DODU DINDON

Cette séquence de lettre est tirée d’un alphabet latin A=

La lettre apparait avec une probabilité avec

* Calculer l’entropie de la séquence X en négligeant l’espace, le trait d’union, l’apostrophe et la ponctuation

1. **Observation d’une densité de probabilité**

Saisir le code de la figure 1 sur l’éditeur de MATLAB (version récente) et procéder à son exécution.

1. Quelle est la fonction réalisée par hist (x,h) à la ligne 10 ?
2. Relever la sortie de ce programme pour L=1, 2, 3 et 4 en définissant les axes.
3. Quelle est la nature de la séquence de symboles générée ?

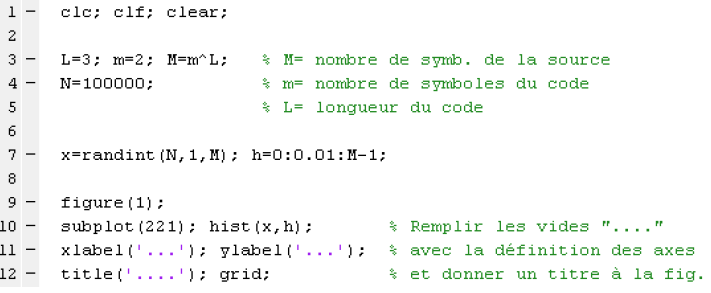


Figure 1 : Génération et observation.

1. **Mesure d’entropie**

Compléter le premier programme avec celui qui est indiqué ci-dessous.

1. Quelle est le contenu de P à l’issu de l’exécution de la figure 2 ?
2. Compléter la ligne 21 pour calculer la quantité d’information associée à chaque symbole (I) ainsi que l’entropie de l’alphabet généré (H).
3. Pourquoi utiliser la base 2 pour le calcul du logarithme ?
4. Relever les valeurs de H pour L=1, 2, 3 et 4. Commentaires.

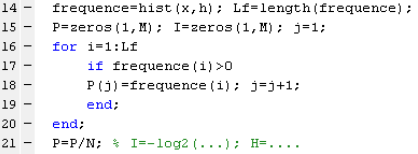


Figure 2 : Génération et observation

1. **Programmation d’une source**

Nous nous intéressons au codage de la source z comme celle programmée par la ligne 29 de la figure 3.

1. Compléter et exécuter ce programme. Comparer P et Pz.
2. Quelle est la nature statistique de la source générée ?
3. Compléter la ligne 43 afin de calculer la quantité d’information associée à chaque symbole (Iz) ainsi que l’entropie de la source (Hz).
4. Comparer l’entropie des deux sources x et z. Commentaires.

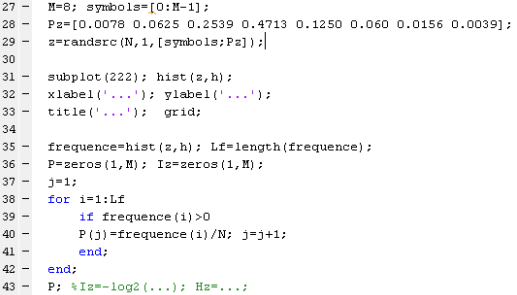


Figure 3 : Programmation d’une source.

1. **Codage et compression**

Le codage et le décodage sont réalisés à l’aide des fonctions huffmanco et huffmandeco de MATLAB comme cela est illustré dans la figure 4.

1. Relever et commenter toutes les sorties de la figure 4 dans la fenêtre de commande.
2. Reporter les symboles, leur code ainsi que la longueur de chaque code dans l’ordre décroissant de la probabilité des symboles. Justifier la valeur de LMoyz.
3. En déduire le degré de compression appliqué à cette source.
4. Construire l’arborescence de Huffman.
5. Justifier le codage obtenu plus haut.

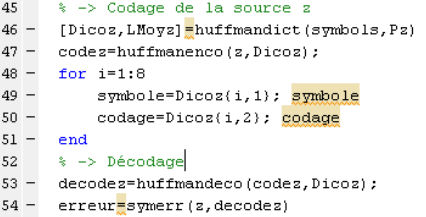


Figure 4 : Codage et décodage

Recommencer l’expérience du codage pour la source x.

1. Quelle est la longueur moyenne du mot-code ?
2. Conclure sur l’intérêt de ce type de codage pour une source équi-répartie.

**Le compte rendu qui doit être déposé une semaine après la date de dépôt du TP, renferme :**

1. **Le fichier.m**
2. **Un document pour répondre aux 20 questions (.pdf)**

***Remarque : Prière de saisir des réponses numérotées et claires.***

***Chargés du module : Mr Abedi et Mr Benbakreti Année : 21/22 Bon courage***