

Exercice académique numéro 3 (à traiter par les candidats des séries autres que la série S)

Compression

Partie A : Code binaire

Dans la mémoire d'un ordinateur, toutes les données sont stockées sous forme numérique en utilisant un code binaire : l'information est chiffrée avec les bits « 0 » et « 1 ». Initialement, le code *ASCII* représentait 128 caractères tous codés avec 7 bits. Un extrait de la table de correspondance de ce code est reproduit en **annexe 3**.

Par exemple : le mot **abel** se code 1100001110001011001011101100 en binaire. Le message **de perga** se code 11001001100101010000011100001100101111001011001111100001.

1. Quel mot se code 11000101101111111001011001011101100 en binaire ?
2. Déterminer le nombre de bits nécessaires pour coder en binaire le message suivant :

ce qui est affirme sans preuve peut etre nie sans preuve

Le but du problème est d'utiliser des tables de correspondance permettant de réduire le nombre de bits nécessaires pour coder numériquement des textes.

Partie B : Code non préfixé

Pour cela, il est possible d'utiliser une méthode statistique reposant sur la fréquence d'apparition des caractères qui le composent : les caractères apparaissant le plus souvent sont codés par un petit nombre de bits alors que les caractères les moins fréquents sont codés par un plus grand nombre de bits.

1. En utilisant la table ci-dessous, coder le message **huffman** .

Caractère	<i>f</i>	<i>a</i>	<i>h</i>	<i>m</i>	<i>n</i>	<i>u</i>
Code compressé	00	10	100	101	110	111

2. Donner deux autres mots pouvant être codés de la même manière.
3. Recopier la table en modifiant une seule case de la ligne « code compressé » pour pouvoir coder et décoder le message **huffman** d'une unique manière.

Partie C : Code préfixé

Dans toute cette partie, on utilise la table de correspondance de l'**annexe 4** pour compresser le message **ce qui est affirme sans preuve peut etre nie sans preuve** .

1. Décoder le code compressé 010010110011001010100.
2. Coder le début de la phrase **ce qui** grâce à cette table.
3. On appelle *taux de compression* le rapport :

$$\text{taux de compression} = \frac{\text{nombre de bits utilisés par le code préfixé}}{\text{nombre de bits utilisés par la code binaire}}$$

- a. Calculer le *taux de compression* pour le message :

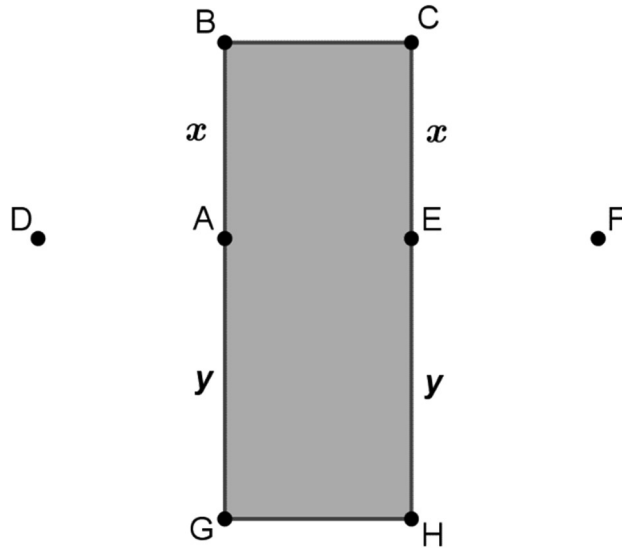
ce qui est affirme sans preuve peut etre nie sans preuve

- b. Interpréter le résultat.

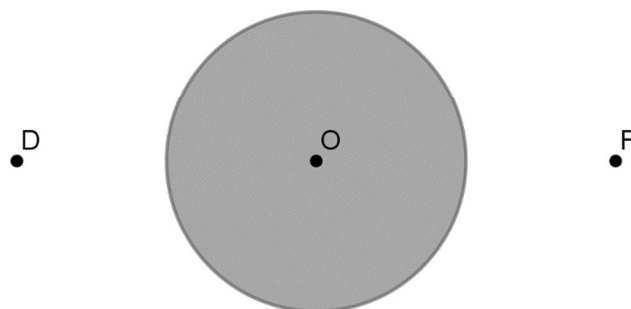
Annexes

Les annexes 1 et 2 sont à COLLER sur la copie

Annexe 1 (Exercice 2 question B.1)



Annexe 2 (Exercice 2 question D.1)



Annexe 3 (Exercice 3 Partie A)

Caractère	Code binaire	Caractère	Code binaire
<i>a</i>	1100001	<i>p</i>	1110000
<i>b</i>	1100010	<i>q</i>	1110001
<i>c</i>	1100011	<i>r</i>	1110010
<i>d</i>	1100100	<i>s</i>	1110011
<i>e</i>	1100101	<i>t</i>	1110100
<i>f</i>	1100110	<i>u</i>	1110101
<i>g</i>	1100111	<i>v</i>	1110110
<i>h</i>	1101000	<i>w</i>	1110111
<i>i</i>	1101001	<i>x</i>	1111000
<i>j</i>	1101010	<i>y</i>	1111001
<i>k</i>	1101011	<i>z</i>	1111010
<i>l</i>	1101100	<i>espace</i>	0100000
<i>m</i>	1101101	<i>virgule</i>	0101100
<i>n</i>	1101110	<i>apostrophe</i>	0100111
<i>o</i>	1101111	<i>point</i>	0101110

Annexe 4 (Exercice 3 Partie C)

Caractère	Code compressé
<i>e</i>	00
<i>espace</i>	111
<i>s</i>	1101
<i>u</i>	1100
<i>r</i>	1011

Caractère	Code compressé
<i>a</i>	1001
<i>t</i>	0101
<i>p</i>	0100
<i>n</i>	0111
<i>i</i>	0110

Caractère	Code compressé
<i>v</i>	10101
<i>f</i>	10100
<i>q</i>	10000
<i>m</i>	100011
<i>c</i>	100010