Problème : On veut relier une source sans mémoire $\boldsymbol{S = \{A,B\}}$ à deux serveurs INTERNET $\boldsymbol{X}$ (entrée) et $\boldsymbol{Y}$ (sortie) grâce à une liaison numérique à haut débit. Le débit symbole d'utilisation de la source $\boldsymbol{S}$ est de $\boldsymbol{D_S = 3\ 10^5\ symb/s}$ et les probabilités des symboles $\boldsymbol{A}$ et $\boldsymbol{B}$ sont $\boldsymbol{p(A) = 0.98 = 1 - p(B)}$. Pour relier ces deux serveurs, on dispose d'un canal binaire symétrique de probabilité d'erreur $\boldsymbol{p = 0{,}02}$.

Nyquist puis Shanonn ont étudié dans la première moitié du $\mathbf{20^e}$ siècle le débit de transmission des canaux bruités. La formule $\boldsymbol{D_{bits/s} = F \times \log_2(1 + \frac{S}{B})}$ donne le débit binaire maximal par unité de temps d'un canal bruité. Dans cette formule, $\boldsymbol{F}$ désigne la fréquence maximale en Hertz utilisable dans le canal, $\boldsymbol{\frac{S}{B}}$ désigne la part de bruit dans un signal reçu. Si par exemple $\boldsymbol{\frac{S}{B} = 1000}$ alors le bruit fait donc varier le signal original d'un millième en moyenne.

Dans la suite , on note $\boldsymbol{D_C}$ le débit binaire par unité de temps du canal symétrique et $\boldsymbol{C}$ la capacité du canal symétrique.

Dans les calculs demandés, on prendra que les 4 chiffres après la virgule

Question préliminaire

1) En sachant que $\boldsymbol{\frac{S}{B} = 1023}$ pour le canal symétrique et sa fréquence maximale autorisée est $\boldsymbol{20000\ Hz}$, calculez le débit maximal par unité de temps théorique $\boldsymbol{D_C}$. **1 pts**
2) Calculer $\boldsymbol{C}$. **1 pt**
3) Calculer l'entropie de la source $\boldsymbol{H(S)}$. **1 pt**
4) Peut-on transmettre le contenu de $\boldsymbol{S}$ via le canal? **2 pts**

Partie A: Codage de source

5) Le codage de Huffman associé à l'extension d'ordre 3 de la source $\boldsymbol{S}$ est donné dans le tableau suivant :

désignation des mots	mots source d'ordre 3	probabilités	mot code
$\boldsymbol{s_0}$	AAA	**0,941**	1
$\boldsymbol{s_1}$	AAB	**0,019**	011
$\boldsymbol{s_2}$	ABA	**0,019**	010
$\boldsymbol{s_3}$	ABB	$\mathbf{3{,}92\ 10^{-4}}$	00011
$\boldsymbol{s_4}$	BAA	**0,019**	001
$\boldsymbol{s_5}$	BAB	$\mathbf{3{,}92\ 10^{-4}}$	00010
$\boldsymbol{s_6}$	BBA	$\mathbf{3{,}92\ 10^{-4}}$	00001
$\boldsymbol{s_7}$	BBB	$\mathbf{8\ 10^{-6}}$	00000

(5.a) Calculer la longueur moyenne $\boldsymbol{\overline{n}}$ correspondant à un symbole source. **1 pt**

(5.b) Soit $\boldsymbol{S'} = \{\mathbf{0,1}\}$ la nouvelle source obtenue à partir de $\boldsymbol{S}$ par le codage de Huffman précédent. Calculer le débit binaire moyen $\boldsymbol{D(S')}$. **1 pt**

(5.c) Quel est le taux de réduction du débit binaire? **1 pt**

Partie B: Codage du canal

6) Afin de réduire le taux d'erreur sur le canal, on se propose de coder le canal en répétant une fois chaque élément binaire d'information. Cela signifie que les mots code sont formés d'un bit d'information et d'un bit de contrôle identique au bit d'information. Les seuls mots code du canal sont donc **00** et **11**. Quel est l'ordre minimum d'extension de la source $\boldsymbol{S}$ nécessaire pour pouvoir transmettre ce code à répétition sur le canal? **4 pts**

7) Au lieu d'utiliser le code à répétition précédent, on se propose maintenant de coder l'information binaire du canal en utilisant le code $\boldsymbol{C(6{,}3)}$ engendré par les équations suivantes :

$$\boldsymbol{c_1 = i_1 + i_2 + i_3}$$
$$\boldsymbol{c_2 = i_1 + i_3}$$
$$\boldsymbol{c_3 = i_2 + i_3}$$

où $\boldsymbol{(i_1,i_2,i_3)}$ sont les bits d'informations et $\boldsymbol{(c_1,c_2,c_3)}$ sont les bits de contrôle. Un mot de code s'écrit donc $\boldsymbol{(i_1,i_2,i_3,c_1,c_2,c_3)}$.

(a) Déterminer une matrice génératrice $\boldsymbol{G}$ de $\boldsymbol{C(6{,}3)}$ sous forme systématique. **1 pt**

(b) Encoder le mot source **011** à l'aide de $\boldsymbol{G}$. **1 pt**

(c) Déterminer la matrice de contrôle $\boldsymbol{H}$. **1 pt**

(d) Calculer la distance minimale du code. **2 pts**

(e) En déduire le nombre d'erreurs que le code peut détecter/corriger. **1 pt**

(f) La table de décodage indiquant les configurations d'erreurs effectivement corrigées de ce code est donnée ci-dessous. Décoder les messages reçus suivants avec la méthode de décodage par syndrome : **2 pts**

$$\boldsymbol{y_1 = 111100, \qquad y_2 = 010011}$$

Syndromes	000	001	011	100	101	110	111
Vecteurs de correction	000000	000001	000010	000100	010000	100000	001000