Codage de l'information

Devoir surveillé n° 2

4 janvier 2017 - Durée 2h - Documents autorisés. Calculatrices **non autorisées**

Veuillez indiquer le numéro de votre groupe de TD sur la copie qu'il est inutile de rendre anonyme, ainsi que votre NIP (figurant sur votre carte d'étudiant).

Ce sujet contient trois exercices indépendants. Prenez 10 mn pour lire l'intégralité du sujet avant de commencer.

Évidemment, toute réponse doit être justifiée.

Exercice 1-1 *Représentation d'information*

Question 1 Donnez la valeur de $\overline{\texttt{A25F}}_{16}$ en binaire puis en décimal.

Question 2

Q 2–1 Expliquez pourquoi, et dans quel cas, les deux nombres -113 et ~~15~~ 143 peuvent être codés par la même chaîne binaire de longueur 8 selon la représentation utilisée.

Q 2–2 De même, quel nombre a la même chaîne binaire de longueur 11 que -224 ?

Question 3 Sachant que UTF-8 permet, en théorie, le codage d'environ 2 millions de caractères, combien d'octets pourraient-être utilisés dans le cadre d'un codage à longueur fixe ?

Question 4 Est-il possible de représenter la valeur 1000,1 en utilisant la représentation IEEE-754 sur 16 bits (5 bits pour l'exposant et 10 bits pour la mantisse) ?

Exercice 1-2 *Codage optimal*

Dans cet exercice, p désigne un nombre réel strictement compris entre 0 et $\frac{1}{2}$, et $q = 1 - p$.

On considère une source de $\mathcal{S} = \{a, b, c, d\}$ de quatre symboles de probabilités respectives p^2, pq, pq et q^2.

Question 1 Exprimez en fonction de p l'entropie de cette source.

Question 2

Q 2–1 Justifiez l'existence d'un codage binaire de cette source ayant pour distribution de longueurs : $[1, 2, 3, 4]$.

Q 2–2 Une telle distribution peut-elle être celle d'un codage optimal ? Si oui donnez un codage pour cette source avec cette distribution de longueurs, dont la longueur moyenne soit la plus faible possible.

Question 3 Même question avec la distribution de longueurs : $[1, 2, 3, 3]$.

Question 4 Donnez les valeurs de p pour lesquelles le codage proposé à la question 3 est effectivement optimal.

Exercice 1-3 *Codes détecteurs/correcteurs*

On considère le codage des mots binaires de longueur 3 en des mots binaires de longueur 15 obtenus en répétant 5 fois le mot à coder.

$$\begin{array}{rccl} \mathbf{c} \ : & \mathbb{F}_2^3 & \longrightarrow & \mathbb{F}_2^{15} \\ & \mathbf{u} & \longmapsto & \mathbf{c}(\mathbf{u}) = \mathbf{u.u.u.u.u} \end{array}$$

On désigne par $C = \mathbf{c}\left(\mathbb{F}_2^3\right)$ le code associé.

On note p la probabilité d'erreur sur un bit dans le canal de transmission supposé symétrique et sans mémoire.

Question 1 Quelle est la distance minimale de C ? Quelles sont les capacités de détection et de correction d'erreurs du codage ?

Question 2 On rappelle que, pour tout $\mathbf{v} \in \mathbb{F}_2^n$ et tout $r \in \mathbb{R}$, $\mathcal{B}(\mathbf{v}, r)$ désigne la boule fermée de centre $\mathbf{v}$ et de rayon r, i.e. l'ensemble des mots de longueur n à distance inférieure ou égale à r.

$$\mathcal{B}(\mathbf{v}, r) = \{\mathbf{v}' \in \mathbb{F}_2^n \mid \mathrm{d}_H(\mathbf{v}, \mathbf{v}') \leq r\}.$$

Soit $\mathbf{u} \in \mathbb{F}_2^3$. Combien y a-t-il de mots de longueur 15 dans $\mathcal{B}(\mathbf{c}(\mathbf{u}), 2)$? Parmi ces mots combien y en a-t-il dans C ?

Question 3 Soit $\mathbf{u} \in \mathbb{F}_2^3$, et $\mathbf{v} = \mathbf{u.u.u.u.u}$ le mot de code associé transmis via le canal. Soit $\mathbf{v}' = \mathbf{v} \oplus \mathbf{e}$ le mot reçu, $\mathbf{e} \in \mathbb{F}_2^{15}$ désignant le mot décrivant les éventuelles erreurs.

Déterminez en fonction de p la probabilité que $\mathbf{v}' \in \mathcal{B}(\mathbf{v}, 2)$, c'est-à-dire que $\mathbf{v}'$ ait été reçu avec deux erreurs ou moins.

Question 4 Pour décoder en cas d'erreur, on décode avec la règle de majorité :

On découpe les 15 bits du mot à décoder en trois listes :
— la liste L_1 des bits de position 1, 4, 7, 10 et 13 ;
— la liste L_2 des bits de position 2, 5, 8, 11 et 14 ;
— et la liste L_3 des bits de position 3, 6, 9, 12 et 15.
Dans chacune de ces listes on choisit le bit majoritaire. Le mot décodé est obtenu en concaténant les trois bits ainsi obtenus.

Q 4–1 Décodez selon cette règle les mots qui suivent :

1. $\mathbf{v}_1 = \texttt{011010011011010}$,
2. $\mathbf{v}_2 = \texttt{101101101101101}$,
3. $\mathbf{v}_3 = \texttt{101101011011110}$.

Q 4–2 Cette règle de décodage permet-elle de décoder correctement certains mots contenant plus d'erreurs que la capacité de correction ?

Question 5 Supposons que l'on définisse un codage $\mathbf{c}'$ utilisant le codage $\mathbf{c}$ et produisant des mots de longueur 20 :

$$\begin{array}{cccc} \mathbf{c}' : & \mathbb{F}_2^3 & \longrightarrow & \mathbb{F}_2^{20} \\ & \mathbf{u} & \longmapsto & \mathbf{c}'(\mathbf{u}) = \mathbf{c}(\mathbf{Parite}(\mathbf{u})) \end{array}$$

sachant que $\mathbf{Parite}(\mathbf{u}) = \mathbf{u} \cdot b$, où b est un bit de parité de manière à ce que le nombre de 1 dans $\mathbf{u} \cdot b$ soit pair.

Q 5–1 Quelle est la distance minimale de $\mathbf{c}'$ ainsi que ses capacités de détection et correction ?

Q 5–2 Quel est le résultat de $\mathbf{c}'(\mathbf{011})$?

Q 5–3 Donnez une matrice génératrice du codage $\mathbf{c}'$.