Examen de THÉORIE DE L'INFORMATION - Ensimag - 1A
Polycopie, notes de cours et calculatrice autorisés

Le sujet est composé de 4 problèmes indépendants. Le barème est donné à titre indicatif.

1 CODAGE SOURCE (4 points)

On considère une source simple S à 5 états notés 'a', 'b', 'c', 'd', 'e' de probabilités $p(a) = 0,04$; $p(b) = 0,06$; $p(c) = 0,4$; $p(d) = 0,2$; $p(e) = 0,3$

1. Quelle est la redondance de cette source S ?
2. Construire C_2 un code de Fano-Shannon pour cette source.
3. Quelle est l'efficacité de C_2 ?
4. Construire C_3 un code de Huffman pour cette source.
5. Quelle est l'efficacité de C_3 ?

2 CODAGE CANAL (6 points)

On considère le code bloc linéaire binaire de matrice génératrice :

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

1. Donner la taille n des mots du code, le nombre k de bits d'information.
2. Dire de combien de mots le code est composé et écrire l'ensemble des mots.
3. Calculer la distance minimale de ce code et en déduire sa capacité de correction d'erreur.
4. Ecrire la matrice de contrôle de parité $\mathbf{H}$.
5. On note $\mathbf{c}$ le mot de code en entrée du canal. Le canal est supposé binaire symétrique de probabilité de transition p. On note $\mathbf{y}$ la séquence associée en sortie du canal.
   (a) En supposant qu'il s'est produit une seule erreur, dire quelles sont les valeurs possibles du syndrome et donner un algorithme de correction d'erreur (autre que la recherche exhaustive).
   (b) Le syndrome vaut 0101, que peut-on dire ?
   (c) Peut-on construire un code à répétition de même rendement et de même distance minimale ? Justifier en cas de réponse négative ou donner sa matrice génératrice dans le cas positif.

3 CANAL À ENTRÉE BINAIRE ET BRUIT TERNAIRE (6 points)

On considère un canal sans mémoire d'entrée $X \in \{-1, +1\}$ binaire et à bruit B additif ternaire à valeurs dans $\{-1, 0; +1\}$ avec les probabilités respectives $\{1/4; 1/2; 1/4\}$. On note $Y = X + B$ la sortie du canal.

1. Donner la matrice de transition de ce canal.
2. Calculer l'entropie conditionnelle $H(Y|X)$.
3. Est-il possible d'atteindre une loi uniforme en sortie de canal ? Dire pourquoi en quelques mots.
4. Quelle est l'entropie maximale de la sortie ? Pour quelle loi d'entrée est-elle atteinte ?
5. En déduire la capacité du canal.

4 OUTILS GÉNÉRAUX - DIVERGENCE DE KULLBACK-LEIBLER [4 points]

Soient 2 sources simples S_A et S_B de même alphabet binaire $\mathcal{A} = \{0; 1\}$ mais de lois de probabilités respectives différentes $P_{S_A} = \{1/2; 1/2\}$ et $P_{S_B} = \{3/4; 1/4\}$.

Un processus aléatoire tire au hasard la source S parmi les 2 possibles de manière équiprobable ($Pr(S = S_A) = Pr(S = S_B) = 1/2$), puis émet une succession indépendante de K variables binaires de la source retenue.

A partir de l'observation (ou réalisation) de la séquence émise $S_{\text{obs}} = [s_1, s_2, \cdots, s_K]$, on doit décider la source émettrice (décision notée $\hat{S}$).

1. On suppose l'observation $S_{\text{obs}} = [0, 1, 0, 0, 1, 0, 1, 0]$ avec $K = 8$.
   La distribution empirique (fréquences) des 0 et 1 est donc $P_{\text{obs}} = \{5/8; 3/8\}$.
   (a) Quelle serait la décision obtenue en utilisant comme critère la distance euclidienne entre les histogrammes empiriques (P_{obs}) et théoriques (P_{S_A} et P_{S_B}) ?
   (b) On rappelle que la règle de décision optimale au sens de la probabilité d'erreur minimale est : on décide $\hat{S} = S_A$ si $Pr(S_{\text{obs}}|S = S_A) > Pr(S_{\text{obs}}|S = S_B)$, et on décide $\hat{S} = S_B$ sinon.
   Pour $S_{\text{obs}} = [0, 1, 0, 0, 1, 0, 1, 0]$, calculer les 2 probabilités $Pr(S_{\text{obs}}|S = S_A)$ et $Pr(S_{\text{obs}}|S = S_B)$ et prendre la décision.
   (c) Démontrer formellement que la règle de décision optimale est équivalente à une règle de comparaison de divergences de Kullback-Leibler à préciser.
   (d) Calculer $D_{KL}(P_{\text{obs}}; P_{S_A})$ et $D_{KL}(P_{\text{obs}}; P_{S_B})$ où D_{KL} est la divergence de Kullback-Leibler. En déduire la décision à prendre.