Examen du cours de Théorie de l'Information et Codage

1. Problème 1: On considère un code de Huffman pour une source U où $\mathbb{P}(u)$ est une puissance de $1/2$ pour tout u.

   (a) Montrer que le code de Huffman aura une longueur moyenne égale à l'entropie de la source. On considère maintenant une source discrète stationnaire: $X_0, X_1, \dots$ c'est à dire on a $\mathbb{P}(X_0) = \mathbb{P}(X_k)$ et $\mathbb{P}(X_1|X_0) = \mathbb{P}(X_{k+1}|X_k)$ pour tout $k \geq 0$ (mais la suite de variables aléatoires n'est pas forcément i.i.d.). Pour simplifier, on supposera que toutes les probabilités considérées sont des puissances de $1/2$. Nous considérons deux méthodes:

   - Méthode 1: on considère les paires successives $(X_1, X_2), (X_3, X_4), \dots$ et on encode chaque paire en utilisant le code de Huffman pour la paire.
   - Méthode 2: On encode chaque X_k en utilisant un code de Huffman basé sur la probabilité conditionelle de X_k sachant la précédente valeur X_{k-1} (Il faut donc utiliser en général $|\mathcal{X}|$ codes de Huffman différents).

   (b) Calculer la longueur moyenne des mots code par lettre pour la méthode 1.

   (c) Calculer la longueur moyenne des mots code par lettre pour la méthode 2.

   (d) Quelle méthode est la meilleure?

2. Problème 2: La course de chevaux. Trois chevaux participent à la course. Si vous pariez 1 euro avant la course, vous en recevez 3 si votre cheval gagne (gain net de 2) et rien s'il n'est pas gagnant (perte de 1). Les probabilités de gagner sont connues et valent $(p_1, p_2, p_3) = (1/2, 1/4, 1/4)$. Vous distribuez la totalité de votre mise sur les chevaux. Soit b_i la fraction correspondant au cheval i. On note $\mathbf{b} = (b_1, b_2, b_3)$, $b_i \geq 0$ et $\sum b_i = 1$. Si le cheval i gagne la course, vous gagnez donc $3b_i$, les autres paris étant perdus. Il y a une infinité de courses (!) dont les résultats sont indépendants les uns des autres et à chaque course, vous misez l'intégralité de votre mise.

   (a) Montrer que si à chaque course, vous jouez toujours le cheval 1 (celui qui a le plus de chance de gagner), vous allez perdre votre mise initiale avec probabilité 1 au bout d'un temps fini.

   (b) On définit

   $$W(\mathbf{b}) = \sum_{i=1}^{3} p_i \log 3b_i.$$

   Maximiser W et en déduire une stratégie (fixe dans le temps) qui permet de gagner de l'argent.

3. Problème 3: quel est le découpage effectué par l'algorithme de Lempel-Ziv pour la suite constante $11111\dots$? Montrer que le nombre de bits codants par symbole pour cette suite tend vers zéro quand $n \to \infty$.

4. Problème 4: On considère le canal binaire sans mémoire à effacement: alphabet d'entrée $A_X = \{0, 1\}$, alphabet de sortie $A_Y = \{0, *, 1\}$, probabilités de transition $p(y|x))$ données par $p(0|0) = p(1|1) = 1 - \alpha$ et $p(*|0) = p(*|1) = \alpha$. On considère que le coût $b(x)$ est donné par $b(0) = 1$ et $b(1) = 2$. Calculer la fonction capacité coût du canal.

5. Problème 5: Calculer la capacité d'une cascade de canaux binaires sans mémoires symmétriques:

$$X_0 \to \boxed{\text{BSC 1}} \to X_1 \to \cdots \to X_{n-1} \to \boxed{\text{BSC n}} \to X_n,$$

chacun avec probabilité d'erreur p. Montrer que si $p \neq 0, 1$, on a $\lim_{n\to\infty} I(X_0, X_n) = 0$.

6. Problème 6: On considère un code binaire ayant M mots code de longueur n. On suppose que ce code peut corriger jusqu'à e erreurs.

   (a) Pour un mot code x, soit S l'ensembles des suites binaires pour lesquelles le décodeur décide x. Montrer que
   $$|S| \geq \sum_{i=0}^{e} \binom{n}{i}.$$

   (b) Montrer que le nombre de mots code M satisfait:
   $$M \leq \frac{2^n}{\sum_{i=0}^{e} \binom{n}{i}}.$$

   (c) Montrer que les codes de Hamming $[n = 2^m - 1, k = n - m, d = 3]$ sont les codes corrigeant une erreur qui pour cette longueur ont le taux le plus élevé possible.

   (d) Un code BCH binaire de longueur $n = 2^m - 1$ et de paramètre $\delta = 2t + 1$ a pour distance minimale d où $d \geq 2t + 1$ est impaire. Montrer que si
   $$\sum_{i=0}^{t+1} \binom{2^m - 1}{i} \geq 2^{mt},$$
   alors $d = 2t + 1$.

7. Problème 7: On considère le code de Hamming étendu: chaque mot code d'un code de Hamming est étendu d'un bit qui vaut 0 si le mot code contient un nombre paire de 1 et 1 sinon. Par exemple le mot code 1000011 devient 10000111. Montrer qu'un tel code a une distance minimale de 4, peut corriger 1 erreur et en détecter 2.

8. Problème 8: Si $n = ab$, montrer que le code binaire BCH de longueur n et paramètre a a pour distance minimale a. On pourra montrer que $1 + x + \cdots + x^{(a-1)b}$ est un mot code.