ENSEIRB **janvier 2007**
Informatique 1ère année

Théorie de l'information

Documents autorisés : une feuille A4 manuscrite recto-verso.
Les exercices sont indépendants. Durée : 1h30

Exercice 1

1. Un code préfixe est-il toujours uniquement déchiffrable ? Un code uniquement déchiffrable est-il toujours préfixe ?

2. La fonction de Kraft d'un code est 0.9. Peut-on en déduire qu'il est uniquement déchiffrable ?

3. La fonction de Kraft d'un code est 1.04. Peut-on en déduire qu'il n'est pas préfixe ? Qu'il n'est pas uniquement déchiffrable ?

4. Le code formé des mots $\{01, 10, 110, 1111, 1110, 001\}$ est-il préfixe ?
Et le code $\{01, 10, 110, 1111, 1110, 0100\}$?

5. Dans une population, 25% des gens sont blonds, 30% ont les yeux bleus, et 75% des blonds ont les yeux bleus. Vous apprenez que quelqu'un a les yeux bleus. Combien d'information avez-vous reçu ? Vous apprenez maintenant que cette même personne a les cheveux blonds. Combien d'information supplémentaire avez-vous reçu ?

Exercice 2

On considère une source constituée par 6 symboles $S = \{x_1, x_2, \ldots, x_6\}$ avec les probabilités $P(x_1) = 1 - t$, $P(x_2) = P(x_3) = \ldots = P(x_6) = \frac{t}{5}$, où t est un réel compris entre 0 et 1.

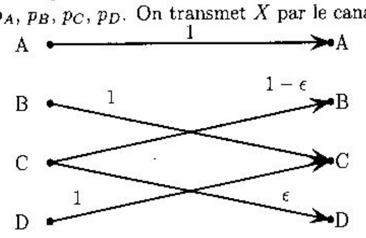
1. Construire le code de Huffmann de S pour $t = 0.5$. Calculer l'entropie de la source, et donner l'efficacité du code.

2. On suppose que $t \leq \frac{5}{6}$.
Donner en fonction de t le code de Huffmann de S (on montrera que ce code a deux formes possibles, suivant que t est inférieur ou supérieur à une valeur t_c à déterminer).

3. On suppose maintenant $t > \frac{5}{6}$. Donner le code de Huffmann de S.

Exercice 3

On dispose d'une source X qui peut envoyer 4 symboles A, B, C, D avec probabilités respectives p_A, p_B, p_C, p_D. On transmet X par le canal ci-dessous, et on note Y le symbole reçu.



1. (a) Déterminer $H(Y)$, $H(X,Y)$, et $I(X,Y)$ en fonction de p_A, p_B, p_C, p_D et ϵ.
   (b) Donner la capacité de ce canal.

2. On décode A par A, B par C, C par B et D par C. Calculer la probabilité d'erreur de décodage d'un symbole en fonction de p_A, p_B, p_C, p_D et ϵ. Pour quelles valeurs des paramètres a-t-on intérêt à décoder C par D ?

Exercice 4

On cherche à coder une source qui envoie les symboles $\{a, b, c, d, e, f, g, h, i\}$ avec les probabilités :

$P(a) = 0.2$	$P(e) = 0.05$
$P(b) = 0.2$	$P(f) = 0.14$
$P(c) = 0.08$	$P(g) = 0.04$
$P(d) = 0.11$	$P(h) = P(i) = 0.09$

1. Calculer l'entropie de la source.
2. Dire dans chacun des cas suivants si un tel code binaire uniquement déchiffrable existe. Justifier une réponse négative, et dans le cas d'une réponse positive, donner un exemple d'un code le plus efficace possible vérifiant la condition, ainsi que son efficacité :
   (a) Un code de longueur constante.
   (b) Un code de longueur moyenne 2.5
   (c) Un code formé avec des mots de longueurs $\{2, 2, 3, 3, 3, 4, 5, 5, 5\}$
   (d) Un code formé avec un mot de longueur 2, les autres mots étant de longueur 3 ou 4.
3. On souhaite qu'un symbole soit codé par un mot de longueur 1, les autres étant tous de même longueur. Quelle longueur faut-il prendre pour que le code soit d'efficacité maximale? Quelle sera son efficacité ?