DÉPARTEMENT "SYSTÈMES INFORMATIQUES FORMELS ET INTELLIGENTS"

THÉORIE DE L'INFORMATION

Devoir surveillé du 30 juin 2008. Rattrapage. *Durée 1h30.*

Document autorisé : *polycopié du cours imprimé uniquement. Les* **calculatrices** *sont admises.*

Exercice 1. Un jeu de cartes contient 4 piques, 2 trèfles, 6 carreaux et 8 coeurs. On tire une carte au hasard. On cherchera ensuite à deviner la couleur de la carte (pique, trèfle, carreau ou coeur) en posant des questions binaires (auxquelles on peut répondre par oui ou par non). L'objectif est de construire la meilleure stratégie de questions.

1. Soit X la variable aléatoire représentant la couleur de la carte. Ses valeurs possibles sont donc $\{\spadesuit, \clubsuit, \diamondsuit, \heartsuit\}$. Donner la distribution de probabilité de X.
2. Calculer l'entropie associée à la distribution de probabilité de X.
3. Quelle est l'entropie associée à une question binaire dont la probabilité de "Oui" est $p \in [0, 1]$? Quelle est sa valeur maximale et pour quelle valeur de p elle est atteinte ?
4. Proposez une stratégie de questions sous forme d'arbre binaire en choisissant à chaque étape la question qui apporte le plus d'information possible selon le contexte.
5. Calculez le nombre moyen de questions selon votre arbre. Comparez le à l'entropie de X. Commentez.

Exercice 2 (Entropie et codage)**.** Soit une source d'alphabet $\Omega = \{x_1, x_2, x_3, x_4, x_5\}$ de distribution de probabilité

X	x_1	x_2	x_3	x_4	x_5
P_X	0.25	0.125	0.125	0.25	0.25

1. Calculer l'entropie de cette source.
2. Montrer qu'il existe un code absolument optimal pour cette source.
3. Construire l'arbre d'un code absolument optimal et en déduire la table de mots de code.

Exercice 3. Soit une source d'alphabet $\Omega = \{a, b, c, d, e\}$ et de distribution de probabilité suivante :

X	a	b	c	d	e
P_X	0.3	0.4	0.05	0.15	0.1

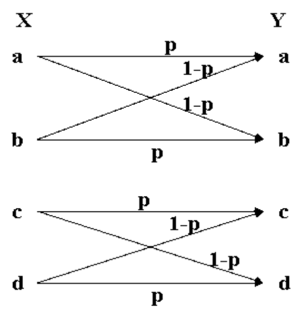
1. Calculer l'entropie de cette source.
2. Quelle est la borne inférieure pour la longueur moyenne d'un code déchiffrable pour cette source ? Existe-t-il un code absolument optimal pour cette source ? Justifier votre réponse.
3. Construire un code de Huffman pour cette source.

4. Calculer la longueur moyenne du code obtenu.

Exercice 4. Soit une source X d'alphabet $\Omega_X = \{a, b, c, d\}$. On suppose que la distribution de probabilité de X est connue :

X	a	b	c	d
P_X	0.2	0.2	0.3	0.3

Les symboles émis sont envoyés via un canal de transmission ayant des perturbations aléatoires. On associe à l'expérience d'observation du symbole reçu la variable aléatoire Y ayant le même alphabet $\Omega_Y = \{a, b, c, d\}$. Les probabilités conditionnelles de réception des différents caractères en fonction des caractères émis sont représentés par le schéma suivant :



Ici $p \in [0, 1]$ est un paramètre. Il représente la probabilité de transmission correcte d'un caractère.

1. A partir du schéma ci-dessus donner la matrice de transition du canal $P(Y|X)$.
2. En utilisant la matrice de transition $P(Y|X)$ et la distribution de probabilité de X, P_X, calculer la distribution de probabilité conjointe $P(X, Y)$. En déduire P_Y, la distribution marginale du récepteur, Y et la distribution conditionnelle $P(X|Y)$.
3. Calculer l'entropie H_X de l'émetteur.
4. Calculer $H(X|Y)$ en fonction du paramètre p.
   Montrer que $H(X|Y) = H_2(p)$ où $H_2(p) = -p\log_2(p) - (1-p)\log_2(1-p)$.
5. Calculer la quantité moyenne d'information transmise $I(X;Y) = H_X - H(X|Y)$ en fonction de p.
   Pour quelle valeur de paramètre $p \in [0, 1]$ la quantité $I(X;Y)$ est **minimale** ?