Théorie de l'information et du codage - Examen écrit
Exercices - 19 janvier 2005

1. **Expérience aléatoire, calculs d'entropies et d'informations mutuelles**

   Soit un jeu de 32 cartes comprenant les valeurs (7, 8, 9, 10, Valet, Dame, Roi, As) pour les quatre couleurs (Coeur, Carreau, Trèfle, Pique). On associe un score à chaque carte selon sa valeur, indépendamment de sa couleur :

Valeur	7	8	9	10	Valet	Dame	Roi	As
Score	0	0	0	2	3	4	10	11

   On tire successivement et sans remise deux cartes de ce jeu. Soient les variables aléatoires :

   - $\mathcal{X}$ désignant la valeur de la première carte (quelle que soit la couleur);
   - $\mathcal{Y}$ désignant la valeur de la deuxième carte (quelle que soit la couleur);
   - $\mathcal{S}$, la somme des scores associés aux deux cartes.

   Calculer :

   (a) $H(\mathcal{X})$, $H(\mathcal{Y})$, $H(\mathcal{Y}|\mathcal{X})$, $H(\mathcal{X}|\mathcal{Y})$, $H(\mathcal{X},\mathcal{Y})$ et $I(\mathcal{X};\mathcal{Y})$.

   (b) $H(\mathcal{S})$, $H(\mathcal{X}|\mathcal{S})$, $H(\mathcal{S},\mathcal{X})$, $H(\mathcal{S}|\mathcal{X})$ et $I(\mathcal{X};\mathcal{S})$.

   (c) $H(\mathcal{S},\mathcal{X},\mathcal{Y})$, $H(\mathcal{X}|\mathcal{Y},\mathcal{S})$, $I(\mathcal{X},\mathcal{Y};\mathcal{S})$ et $I(\mathcal{X};\mathcal{Y}|\mathcal{S})$.

2. **Réseaux bayesiens et modélisation de sources d'information**

   Soient trois suites de variables aléatoires binaires

$$\begin{array}{l}\mathcal{A}_0, \mathcal{A}_1, \ldots \\ \mathcal{B}_0, \mathcal{B}_1, \ldots \\ \mathcal{C}_0, \mathcal{C}_1, \ldots\end{array}$$

   et une distribution conjointe $P(\mathcal{A}_0, \mathcal{B}_0, \mathcal{C}_0, \mathcal{A}_1, \mathcal{B}_1, \mathcal{C}_1, \ldots)$ telle que

$$\begin{aligned} P(\mathcal{A}_0, \mathcal{B}_0, \mathcal{C}_0) &= P(\mathcal{A}_0)P(\mathcal{B}_0|\mathcal{A}_0)P(\mathcal{C}_0|\mathcal{B}_0) \\ P(\mathcal{A}_i, \mathcal{B}_i, \mathcal{C}_i|\mathcal{A}_0, \mathcal{B}_0, \mathcal{C}_0, \ldots, \mathcal{A}_{i-1}, \mathcal{B}_{i-1}, \mathcal{C}_{i-1}) &= P(\mathcal{A}_i)P(\mathcal{B}_i|\mathcal{A}_i, \mathcal{B}_{i-1})P(\mathcal{C}_i|\mathcal{B}_i) \text{ pour } i \geq 1 \end{aligned}$$

   Trois sources S_1, S_2 et S_3 sont modélisées par, respectivement, les suites de variables aléatoires

$$\begin{array}{l}\mathcal{A}_0, \mathcal{A}_1, \ldots \\ \mathcal{C}_0, \mathcal{C}_1, \ldots \\ \mathcal{X}_0, \mathcal{X}_1, \ldots\end{array}$$

   avec $\mathcal{X}_i = \{\mathcal{B}_i, \mathcal{C}_i\}$.

(a) Dessiner le graphe d'un réseau bayesien sur les variables $\{\mathcal{A}_0, \mathcal{B}_0, \mathcal{C}_0, \mathcal{A}_1, \mathcal{B}_1, \mathcal{C}_1, \dots\}$ représentant les indépendances impliquées par la factorisation de $P(\mathcal{A}_0, \mathcal{B}_0, \mathcal{C}_0, \mathcal{A}_1, \mathcal{B}_1, \mathcal{C}_1, \dots)$. Représenter au minimum les neuf premières variables.

(b) En vous aidant éventuellement du graphe du point 2a, discuter, en fonction des indépendances qui peuvent être vérifiées pour la loi $P(\mathcal{A}_0, \mathcal{B}_0, \mathcal{C}_0, \mathcal{A}_1, \mathcal{B}_1, \mathcal{C}_1, \dots)$, la possibilité pour les sources S_1, S_2 et S_3 d'être sans mémoire ou de Markov.

Supposons que

$$\begin{aligned} P(\mathcal{A}_0 = 0) &= 0.2 \\ P(\mathcal{B}_0 = 0|\mathcal{A}_0 = 0) &= 0.5 \\ P(\mathcal{B}_0 = 0|\mathcal{A}_0 = 1) &= \frac{1}{12} \\ P(\mathcal{C}_0 = 0|\mathcal{B}_0 = 0) &= 0.3 \\ P(\mathcal{C}_0 = 0|\mathcal{B}_0 = 1) &= 0.78 \end{aligned}$$

et que, pour $i \geq 1$,

$$\begin{aligned} P(\mathcal{A}_i = 0) &= 0.2 \\ P(\mathcal{B}_i = 0|\mathcal{A}_i = 0, \mathcal{B}_{i-1} = 0) &= 1 \\ P(\mathcal{B}_i = 0|\mathcal{A}_i = 0, \mathcal{B}_{i-1} = 1) &= 0.4 \\ P(\mathcal{B}_i = 0|\mathcal{A}_i = 1, \mathcal{B}_{i-1} = 0) &= 0.5 \\ P(\mathcal{B}_i = 0|\mathcal{A}_i = 1, \mathcal{B}_{i-1} = 1) &= 0 \\ P(\mathcal{C}_i = 0|\mathcal{B}_i = 0) &= 0.4 \\ P(\mathcal{C}_i = 0|\mathcal{B}_i = 1) &= 0.4 \end{aligned}$$

(c) Les sources S_1, S_2 et S_3 sont-elles sans mémoire, de Markov?

(d) Dessiner le diagramme d'état de chaque source sans mémoire ou de Markov.

(e) S_1, S_2 et S_3 sont-elles stationnaires et ergodiques? Expliquer.

(f) Calculer l'entropie (ou débit d'entropie) de chacune des trois sources S_1, S_2 et S_3.

3. **Propriétés de processus**

On dispose d'une pièce possiblement biaisée. Désignons par p la probabilité de tomber sur pile.

(a) On lance la pièce et on note par $\mathcal{X}$ la variable aléatoire qui désigne le résultat de cette expérience. Quelle est l'entropie de $\mathcal{X}$? Quelle est la valeur maximale de cette entropie et la valeur de p qui réalise ce maximum ?

On répète cette expérience indéfiniment et on déduit de la suite $\mathcal{X}_i$ ainsi générée une suite de variables $\mathcal{Z}_i$ de la manière suivante :

- initialisation : $\mathcal{Z}_0 = 0$; i:=1
- boucle :
  - si $\mathcal{X}_i = pile$ alors $\mathcal{Z}_i = \mathcal{Z}_{i-1}$
  - si $\mathcal{X}_i = face$ alors $\mathcal{Z}_i = \mathcal{Z}_{i-1} + 1$
  - $i := i + 1$

(b) Quelles sont les valeurs possibles de $\mathcal{Z}_i$

(c) $\mathcal{Z}$ est-il un processus stationnaire ? ergodique ? de Markov ? sans mémoire ?

Si $p = 0.3$,

(d) calculer l'entropie des 100 premiers symboles générés par ce processus.

(e) Quel est le débit d'entropie de ce processus ?

(f) Quel est le nombre moyen de lancers nécessaires pour que le processus $\mathcal{Z}$ prenne pour la première fois la valeur m ?

4. **Codes de Huffman**

(a) Soit une source stationnaire et sans mémoire $\mathcal{S}$ d'alphabet binaire $\mathcal{S} = \{S_1, S_2\}$ avec $P(S_1) = 0.7$.

Construisez un code de Huffman binaire pour les extensions $\mathcal{S}^2$ et $\mathcal{S}^3$ de cette source et comparez les longueurs moyennes de ces codes avec les entropies $H(\mathcal{S}^2)$ et $H(\mathcal{S}^3)$.

(b) Considérez une source d'alphabet quaternaire quelconque sans mémoire. Trouvez une distribution de probabilités (p_1, p_2, p_3, p_4) telle qu'il existe deux codes de Huffman binaires pour cette distribution qui donnent lieu à deux ensembles différents de longueurs de mots.

(c) Soient $\ell_1, \ell_2, \ldots, \ell_{10}$ les longueurs des mots de code d'un code de Huffman binaire correspondant à une distribution de probabilités $(p_1, p_2, \ldots, p_{10})$ telle que $p_1 \geq p_2 \geq \ldots \geq p_{10}$. Supposons que l'on dérive une nouvelle distribution de probabilités en décomposant p_{10} comme suit :

$$(p_1^{'}, p_2^{'}, \ldots, p_{10}^{'}, p_{11}^{'}) = (p_1, p_2, \ldots, p_9, \alpha p_{10}, (1-\alpha)p_{10}) \qquad 0 \leq \alpha \leq 1$$

Déterminez, en fonction de $\ell_1, \ell_2, \ldots, \ell_{10}$, les longueurs des mots de code du code de Huffman construit pour cette nouvelle distribution de probabilités.

Formules utiles

Entropie, information, divergence

$$H(\mathcal{X}) \triangleq -\sum_{i=1}^{n} P(X_i)\log P(X_i) \tag{F.1}$$

$$H(\mathcal{X}_1,\mathcal{X}_2,\ldots,\mathcal{X}_m) \triangleq -\sum_{i_1=1}^{n_1}\cdots\sum_{i_m=1}^{n_m} P(X_{1,i_1},\ldots,X_{m,i_m})\log P(X_{1,i_1},\ldots,X_{m,i_m}) \tag{F.2}$$

$$H(\mathcal{X}|Y_j) = -\sum_{i=1}^{n} P(X_i|Y_j)\log P(X_i|Y_j) \tag{F.3}$$

$$H(\mathcal{X}|\mathcal{Y}) \triangleq -\sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i,Y_j)\log P(X_i|Y_j) \tag{F.4}$$

$$H(\mathcal{X}|\mathcal{Y}) = \sum_{j=1}^{m} P(Y_j)H(\mathcal{X}|Y_j) \tag{F.5}$$

$$I(\mathcal{X};\mathcal{Y}) \triangleq \sum_{i=1}^{n}\sum_{j=1}^{m} P(X_i,Y_j)\log\frac{P(X_i,Y_j)}{P(X_i)P(Y_j)} \tag{F.6}$$

$$I(\mathcal{X};\mathcal{Y}) = H(\mathcal{X}) - H(\mathcal{X}|\mathcal{Y}) \tag{F.7}$$

$$= H(\mathcal{Y}) - H(\mathcal{Y}|\mathcal{X}) = H(\mathcal{X}) + H(\mathcal{Y}) - H(\mathcal{X},\mathcal{Y}) \tag{F.8}$$

$$I(\mathcal{X};\mathcal{Y}|\mathcal{Z}) \triangleq H(\mathcal{X}|\mathcal{Z}) - H(\mathcal{X}|\mathcal{Y},\mathcal{Z}) \tag{F.9}$$

$$I(\mathcal{X};\mathcal{Y}|\mathcal{Z}) = \sum_{i,j,k} P(X_i,Y_j,Z_k)\log\frac{P(X_i,Y_j|Z_k)}{P(X_i|Z_k)P(Y_j|Z_k)} \tag{F.10}$$

$$D(P||Q) \triangleq \sum_{\omega\in\Omega} P(\omega)\log\frac{P(\omega)}{Q(\omega)} \tag{F.11}$$

Additivité et chaînage

$$H(\mathcal{X},\mathcal{Y}) = H(\mathcal{X}) + H(\mathcal{Y}|\mathcal{X}) = H(\mathcal{Y}) + H(\mathcal{X}|\mathcal{Y}) \tag{F.12}$$

$$H(\mathcal{X},\mathcal{Y}|\mathcal{Z}) = H(\mathcal{X}|\mathcal{Z}) + H(\mathcal{Y}|\mathcal{X},\mathcal{Z}) \tag{F.13}$$

$$H(\mathcal{X}_1,\mathcal{X}_2,\ldots,\mathcal{X}_n) = H(\mathcal{X}_1) + H(\mathcal{X}_2|\mathcal{X}_1) + H(\mathcal{X}_3|\mathcal{X}_2,\mathcal{X}_1) + \ldots + H(\mathcal{X}_n|\mathcal{X}_{n-1},\ldots,\mathcal{X}_1)$$

$$\triangleq \sum_{i=1}^{n} H(\mathcal{X}_i|\mathcal{X}_{i-1},\ldots,\mathcal{X}_1) \tag{F.14}$$

$$H(\mathcal{X}_1,\mathcal{X}_2,\ldots,\mathcal{X}_n|\mathcal{Y}) = H(\mathcal{X}_1|\mathcal{Y}) + H(\mathcal{X}_2|\mathcal{X}_1,\mathcal{Y}) + \ldots + H(\mathcal{X}_n|\mathcal{X}_{n-1},\ldots,\mathcal{X}_1,\mathcal{Y})$$

$$\triangleq \sum_{i=1}^{n} H(\mathcal{X}_i|\mathcal{X}_{i-1},\ldots,\mathcal{X}_1,\mathcal{Y}) \tag{F.15}$$

$$I(\mathcal{X}_1,\mathcal{X}_2,\ldots,\mathcal{X}_n;\mathcal{Y}) = \sum_{i=1}^{n} I(\mathcal{X}_i;\mathcal{Y}|\mathcal{X}_{i-1},\ldots,\mathcal{X}_1) \tag{F.16}$$

Positivité, convexité et monotonicité

$$H(\mathcal{X}) \geq 0 \quad \text{et} \quad H(\mathcal{X},\mathcal{Y}) \geq 0 \quad \text{et} \quad H(\mathcal{X}|\mathcal{Y}) \geq 0 \tag{F.17}$$

$$H(\mathcal{X},\mathcal{Y}) \leq H(\mathcal{X}) + H(\mathcal{Y}) \leq 2H(\mathcal{X},\mathcal{Y}) \tag{F.18}$$

$$0 \leq H(\mathcal{X}) \leq H(\mathcal{X},\mathcal{Y}) \leq H(\mathcal{X},\mathcal{Y},\mathcal{Z}) \leq \ldots \tag{F.19}$$

$$H(\mathcal{X}) \geq H(\mathcal{X}|\mathcal{Y}) \geq H(\mathcal{X}|\mathcal{Y},\mathcal{Z}) \geq \ldots \geq 0 \tag{F.20}$$

$$I(\mathcal{X};\mathcal{Y}) \geq 0 \quad \text{et} \quad I(\mathcal{X};\mathcal{Y}|\mathcal{Z}) \geq 0 \tag{F.21}$$

$$0 \leq I(\mathcal{X};\mathcal{Y}) \leq I(\mathcal{X};\mathcal{Y},\mathcal{Z}) \leq I(\mathcal{X},\mathcal{T};\mathcal{Y},\mathcal{Z}) \tag{F.22}$$

$$D(P||Q) \geq 0 \tag{F.23}$$

Non-création d'information
Si $\mathcal{X} \leftrightarrow \mathcal{Y} \leftrightarrow \mathcal{Z}$ forment une chaîne de Markov alors

$$I(\mathcal{X};\mathcal{Y}) \geq I(\mathcal{X};\mathcal{Z}) \quad \textit{et} \quad I(\mathcal{Y};\mathcal{Z}) \geq I(\mathcal{X};\mathcal{Z}) \tag{F.24}$$

En particulier $\mathcal{Z} = g(\mathcal{Y})$

$$I(\mathcal{X};\mathcal{Y}) \geq I(\mathcal{X};g(\mathcal{Y})) \tag{F.25}$$

Inégalité de Kraft
Soient $n_1, \ldots, n_Q$ des longueurs de mots candidates pour coder une source Q-aire dans un alphabet q-aire. Alors l'inégalité *de Kraft*

$$\sum_{i=1}^{Q} q^{-n_i} \leq 1 \tag{F.26}$$

est une condition nécessaire et suffisante d'existence d'un code déchiffrable respectant ces longueurs de mots.