Master CSI 1 2018-2019

Théorie de l'information : Examen du 19 décembre 2018

Master Sciences et Technologies, mention Mathématiques ou Informatique, parcours Cryptologie et Sécurité informatique

Responsable : Gilles Zémor

Durée : 3h. Sans document. Les exercices sont indépendants.

– EXERCICE 1. On forme le quintuplet X_1, X_2, X_3, X_4, X_5 de variables aléatoires à valeurs dans $\{1, 2, 3, 4, 5\}$ en choisissant au hasard avec loi uniforme une ligne de la matrice :

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 4 & 5 \\ 3 & 2 & 1 & 4 & 5 \\ 4 & 3 & 2 & 1 & 5 \\ 5 & 4 & 3 & 5 & 1 \end{bmatrix}.$$

Calculer $H(X_i)$ pour $i = 1 \ldots 5$ et $H(X_{i+1}|X_i)$ pour $i = 1 \ldots 4$.

– EXERCICE 2. Soient X_0 et X_1 deux variables aléatoires. On tire à pile ou face pour produire la variable Z, $P(Z = 0) = P(Z = 1) = 1/2$. On définit ensuite la variable $Y = X_Z$, c'est-à-dire la variable qui vaut X_0 si $Z = 0$ et X_1 si $Z = 1$.

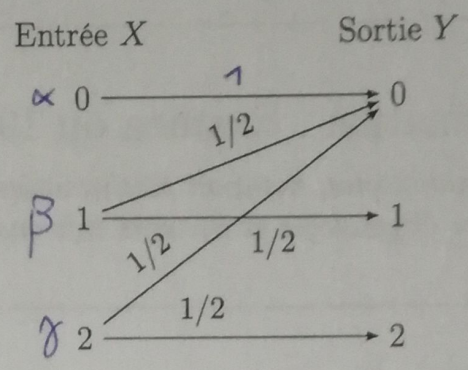
a) Calculer de deux manières $H(Z, Y)$ pour en déduire que

$$H(Y) \geqslant \frac{1}{2}(H(X_0) + H(X_1)).$$

b) Montrer que s'il y a égalité dans l'inégalité ci-dessus, alors X_0 et X_1 prennent leur valeurs dans le même ensemble et ont la même loi.

– EXERCICE 3. Est-ce qu'un code linéaire ternaire (sur l'alphabet $\{0, 1, -1\}$) de paramètres $[12, 7, 6]$ existe ? On pourra se ramener au cas d'une distance minimale impaire.

– EXERCICE 4. On considère le canal représenté par la figure suivante :



où $P(Y = 0|X = 1) = P(Y = 0|X = 2) = 1/2$. On pose $\alpha = P(X = 0)$, $\beta = P(X = 1)$ et $\gamma = P(X = 2)$.

a) Soit Z la variable de Bernoulli qui vaut 0 lorsque $Y = 0$ et 1 sinon. Montrer que $H(Y) = H(Z) + H(Y|Z)$ et en déduire que lorsque l'information mutuelle $I(X, Y)$ est maximale, alors on a $\beta = \gamma = (1 - \alpha)/2$.

b) On suppose dorénavant $\beta = \gamma$. Montrer que

$$I(X, Y) = \frac{1+\alpha}{2} \log_2 \frac{2}{1+\alpha} + \frac{1-\alpha}{2} \log_2 \frac{1}{1-\alpha}.$$

c) En déduire que $I(X, Y)$ est maximale pour $\alpha = 1/3$ et donner la capacité du canal.

– EXERCICE 5. Alice souhaite communiquer un message secret $\mathbf{s} \in \{0, 1\}^3$ à Bob. Pour cela Alice et Bob conviennent que Alice va envoyer à Bob un vecteur binaire $\mathbf{x} \in \mathbb{F}_2^7$ tel que $\sigma(\mathbf{x}) = \mathbf{H}\mathbf{x}^T = \mathbf{s}$, où $\mathbf{H}$ est la matrice de parité d'un code de Hamming $[7, 4, 3]$. Alice choisit $\mathbf{x}$ aléatoirement avec une loi uniforme parmi tous les vecteurs de syndrome $\sigma(\mathbf{x}) = \mathbf{s}$. Puis Alice envoit $\mathbf{x}$ à Bob sur un certain canal de transmission. Bob obtient $\mathbf{x}$ sans erreur et peut reconstituer $\mathbf{s} = \sigma(\mathbf{x})$. Un espion qui écoute la transmission obtient une version bruitée de $\mathbf{x}$. Très précisément, l'espion obtient $\mathbf{y} = \mathbf{x} + \mathbf{e}$ où $\mathbf{e}$ est un vecteur aléatoire choisi avec une loi uniforme dans l'ensemble à huit éléments constitué du vecteur nul et des sept mots de poids 1. Le but de l'exercice est de montrer que l'espion n'obtient aucune information sur $\mathbf{s}$, c'est-à-dire que $H(\mathbf{s}|\mathbf{y}) = H(\mathbf{s})$. Il y a autant d'incertitude sur $\mathbf{s}$ avec ou sans la connaissance de $\mathbf{y}$.

a) Que vaut $H(\mathbf{x}|\mathbf{s})$?

b) Montrer que si on connaît $\mathbf{s}$ et $\mathbf{y}$ alors on connaît $\mathbf{x}$ et $\mathbf{e}$. En déduire $H(\mathbf{s}, \mathbf{y}) = H(\mathbf{x}, \mathbf{e}) = H(\mathbf{x}) + H(\mathbf{e})$.

c) Montrer que $H(\mathbf{x}) = H(\mathbf{x}, \mathbf{s}) = H(\mathbf{s}) + 4$.

d) Que vaut $H(\mathbf{e})$? En déduire que $H(\mathbf{s}|\mathbf{y}) = H(\mathbf{s})$.

– EXERCICE 6. On considère le code linéaire binaire C donné par la matrice de parité

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

a) Trouver une matrice de parité de C sous forme systématique $[A|I_5]$.

b) Que vaut la distance minimale d de C ?

c) Combien y a-t-il de mots de poids d dans C ?

d) Quel est le plus grand entier e tel que n'importe quelle configuration de e effacements peut être corrigée ?

e) Quel est le plus petit entier E tel qu'aucune configuration de E effacements ne peut être corrigée ?

f) Un mot $\mathbf{x}$ du code C est corrompu par une erreur et deux effacements pour donner le 10-uple

$$\mathbf{y} = [??10010000].$$

Que vaut $\mathbf{x}$?

g) On appelle *rayon de recouvrement* du code C le plus petit entier t tel que pour tout $\mathbf{y} \in \mathbb{F}_2^{10}$, il existe un mot de code $\mathbf{x} \in C$ avec $d(\mathbf{x}, \mathbf{y}) \leqslant t$. Que vaut le rayon de recouvrement t de C ?

h) Soit $\mathbf{y}$ un mot de poids t, autrement dit à distance t du mot de code 0. Montrer que le nombre de mots $\mathbf{x}$ *non nuls* de C tels que $d(\mathbf{x}, \mathbf{y}) = t$ vaut :
— soit 1,
— soit 4.
Combien y a-t-il de mots de poids t correspondant à chacun des de ces deux cas ?