Master CSI 1 2022-2023

Théorie de l'information : Examen du 9 mai 2023

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 tire à pile ou face quatre fois de suite. On note X le nombre maximum de «face» consécutifs et Y le nombre maximum de «pile» consécutifs. Par exemple si le résultat des lancers est $FFPP$ on a $X = 2$ et $Y = 2$. Si le résultat est $FPFF$ on a $X = 2$ et $Y = 1$.

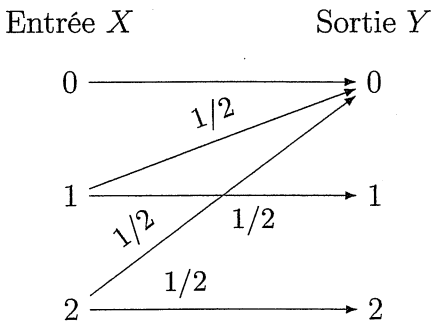
Calculer l'information mutuelle entre X et Y.

– EXERCICE 2. Un joueur A tire à pile ou face cinq fois de suite. On note X le nombre de fois où sort «face».

a) Construire un code de Huffman pour X.

b) Un joueur B doit découvrir la valeur de X en posant à A des questions dont la réponse est «oui» ou «non». Une procédure est dite optimale si elle permet au joueur B de poser une suite de questions successives dont les réponses déterminent X, et telle que le nombre moyen de questions est minimum.
- Quel est le nombre moyen de questions pour une procédure optimale ?
- Quelle est la première question de la procédure optimale ?

– EXERCICE 3. On considère le canal représenté par la figure suivante :



où $P(Y=0|X=1) = P(Y=0|X=2) = 1/2$.

On suppose que la loi de l'entrée X est donnée par $P(X=0) = 1/3, P(X=1) = a, P(X=2) = b$, où $a+b = 2/3$.

a) Donner une expression de l'information mutuelle $I(X,Y)$.

b) Sous l'hypothèse $P(X=0) = 1/3$, en déduire que $I(X,Y)$ est maximale pour $a = b$.

c) En admettant que le maximum de $I(X,Y)$ est atteint pour $P(X=0) = 1/3$, en déduire la capacité du canal, que l'on exprimera sous la forme $\log_2 x$, où x est un nombre rationnel simple.

– EXERCICE 4. Soit $\mathbf{H}$ une matrice de parité fixée d'un code de Hamming $[7,4,3]$.

Alice souhaite communiquer un message secret $\mathbf{s} \in \{0,1\}^3$ à Bob. Pour cela Alice et Bob conviennent du protocole suivant. Alice envoie à Bob sur un certain canal de transmission un vecteur binaire $\mathbf{x} \in \mathbb{F}_2^7$ tel que $\sigma(\mathbf{x}) = \mathbf{H}\mathbf{x}^\intercal = \mathbf{s}$.

Le vecteur $\mathbf{x}$ choisi par Alice est aléatoire avec une loi uniforme parmi tous les vecteurs de syndrome $\sigma(\mathbf{x}) = \mathbf{s}$. Le canal est tel que Bob obtient $\mathbf{x}$ sans erreur et peut donc reconstituer $\mathbf{s} = \sigma(\mathbf{x})$. Par contre, 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 suivant 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 $H(\mathbf{s}|\mathbf{y}) = H(\mathbf{s})$. L'espion a autant d'incertitude sur $\mathbf{s}$ avec ou sans la connaissance de $\mathbf{y}$.

Attention, on ne suppose rien a priori sur la loi de $\mathbf{s}$.

a) Pour un $\mathbf{s}$ fixé, combien de valeurs peut prendre $\mathbf{x}$? Et donc 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 la loi de $\mathbf{y}$ ne peut qu'être uniforme et qu'on a $H(\mathbf{s}|\mathbf{y}) = H(\mathbf{s})$.

– EXERCICE 5. Quelles conditions la matrice de parité d'un code linéaire ternaire (d'alphabet $\mathbb{Z}/3\mathbb{Z}$) doit elle vérifier pour que le code ait une distance minimale au moins 3 ? Quelle est la longueur maximale d'un code ternaire de distance minimale au moins 3, dont une matrice de parité a trois lignes ?

– EXERCICE 6. Soit C un code paramètres $[n, k, d]$. On définit le code *poinçonné* $\mathrm{Poinc}(C)$ de C comme l'ensemble des $(n-1)$-uples $(x_1, x_2, \ldots, x_{n-1})$ pour lesquels il existe un n-uple $(x_1, \ldots, x_{n-1}, x_n)$ appartenant à C. En d'autres termes, $\mathrm{Poinc}(C)$ est l'ensemble des mots obtenus à partir des mots de C en oubliant la dernière coordonnée.

On définit également le code *raccourci* $\mathrm{Rac}(C)$ de C comme étant l'ensemble des $(n-1)$-uples $(x_1, x_2, \ldots, x_{n-1})$ tels que $(x_1, x_2, \ldots, x_{n-1}, 0) \in C$.

a) Que pouvez-vous dire des dimensions et distances minimales des codes $\mathrm{Poinc}(C)$ et $\mathrm{Rac}(C)$?

b) Montrer que $\mathrm{Poinc}(C)^{\perp} = \mathrm{Rac}(C^{\perp})$.

– EXERCICE 7. On considère le code C dont une matrice de parité $\mathbf{H}$ est donnée par

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

a) En remarquant que les cinq premières colonnes de $\mathbf{H}$ sont des décalées circulaires de la première, et que les cinq suivantes sont des décalées circulaires de la sixième, montrer que si un mot $[x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}]$ est dans C, alors le mot $[x_5, x_1, x_2, x_3, x_4, x_{10}, x_6, x_7, x_8, x_9]$ l'est également.

b) Donner, en le justifiant, les paramètres du code C.

c) Donner, sans utiliser de pivot de Gauss fastidieux, la matrice génératrice de C qui est de la forme $[I \mid A]$ où I est la matrice identité 5×5 et A est une matrice 5×5.

d) Quels sont les paramètres du code dual de C ?

e) Le mot [?111011000] est un mot $\mathbf{c}$ de C qui a subi une erreur et un effacement. Trouver $\mathbf{c}$.