Master CSI 1 — 2019-2020

Théorie de l'information : DS du 6 novembre 2019

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

Responsable : Gilles Zémor

Durée : 1h30. Sans document. Les exercices sont indépendants.

– EXERCICE 1. On considère deux variables Y_1 et Y_2, chacune prenant ses valeurs dans l'ensemble $\{01, 10\}$ avec la loi uniforme. On suppose Y_1 et Y_2 indépendantes. On écrira $Y_1Y_2 = X_1X_2X_3X_4$ où les X_i sont des bits aléatoires, et on note $X = X_1X_2X_3X_4$.

a) Calculer $H(X)$, $H(X_2X_3)$ et $H(X_2X_3|X_1)$.

b) Soit σ une permutation aléatoire de l'ensemble $\{1, 2, 3, 4\}$, choisie indépendamment de X et uniformément. On forme le quadruplet

$$X^\sigma = X_{\sigma(1)}X_{\sigma(2)}X_{\sigma(3)}X_{\sigma(4)}.$$

Que vaut $H(X^\sigma)$?

c) Calculer $H(X^\sigma|\sigma)$, $H(X|X^\sigma)$, $H(\sigma|X^\sigma)$.

– EXERCICE 2. Soient X et Y deux variables aléatoires. On pose $Z = X + Y$.

a) Montrer que $H(Z|X) = H(Y|X)$.

b) Montrer que lorsque X et Y sont indépendantes alors $H(X) \leqslant H(Z)$ et $H(Y) \leqslant H(Z)$.

c) Trouver X et Y tels que $H(X) > H(Z)$ et $H(Y) > H(Z)$.

d) Quand a-t-on $H(Z) = H(X) + H(Y)$?

– EXERCICE 3. Soient X et Y deux variables aléatoires, et soit f une fonction définie sur l'ensemble des valeurs prises par Y.

Montrer que

$$H(X|f(Y)) + H(Y|X, f(Y)) = H(Y|f(Y)) + H(X|Y)$$

et en déduire que $H(X|f(Y)) \geqslant H(X|Y)$.

– EXERCICE 4. On considère une variable aléatoire X prenant ses valeurs dans l'ensemble $\{1, 2, 3, 4, 5, 6, 7\}$, de loi de probabilité associée

$$p = \left(\frac{7}{30}, \frac{7}{30}, \frac{1}{5}, \frac{2}{15}, \frac{1}{10}, \frac{1}{15}, \frac{1}{30}\right).$$

Quelles sont les distributions des longueurs $\ell_1, \ell_2, \ldots, \ell_7$ possibles pour les codes de Huffman associés à cette loi ? Donner un code de Huffman correspondant dans chacun des cas. Quelle est la longueur moyenne de ces codes ?

– EXERCICE 5. Soit p une loi de probabilité dont la plus grande probabilité est p_1. Montrer que si $p_1 < 1/3$ alors un code de Huffman n'encode jamais le symbole de probabilité p_1 par un mot de longueur 1.

– EXERCICE 6. Soit une variable aléatoire X prenant m valeurs avec une loi $p = (p_1, \ldots, p_m)$. On code cette variable avec un code C préfixe de distribution des longueurs $\ell_1, \ell_2, \ldots, \ell_m$. On suppose $\sum_{i=1}^{m} 2^{-\ell_i} = 1$ et on pose $q = (q_1, q_2, \ldots, q_m)$ avec $q_i = 2^{-\ell_i}$. Montrer que la longueur moyenne du codage de X par le code C vaut :

$$\overline{\ell} = H(p) + \mathrm{D}(p \,||\, q).$$