Master CSI 1 — 2024-2025

Théorie de l'information, 4TCY806U : DSI du 25 février 2025

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

Responsable : Elena Berardini

Durée : 1h30. Sans document. Les exercices sont indépendants. Toutes les réponses doivent être justifiées.

– EXERCICE 1. **Entropie.**

Dans cet exercice $X : \Omega \to \mathcal{X}$ est une variable aléatoire.

1 **a)** Supposons $\mid \mathcal{X} \mid = 8$ et que X suit la loi uniforme. Justifier que $H(X) = 3$.

1 **b)** Montrer un exemple d'une variable aléatoire X qui a plus de 8 elements, ne suit pas la loi uniforme, et satisfait $H(X) = 3$.

Soit maintenant $g : \mathcal{X} \to \mathcal{Y}$ une fonction. On définit la variable aléatoire $Y : \Omega \to \mathcal{Y}$ par $Y(\omega) = g(X(\omega))$. Le *principe de non-création d'information* affirme alors que $H(Y \mid X) = 0$.

2 **c)** Montrer que $H(X) \geqslant H(g(X))$ (*Conseil :* utiliser plusieurs expressions de $H(X, g(X))$).

1 **d)** Montrer qu'on a toujours $H(X) = H(2X)$.

1 **e)** A t-on toujours $H(X) = H(X^2)$? Justifier.

3 **f)** Prouver le principe de non-création d'information.

– EXERCICE 2. **Information mutuelle.**

On jette deux dés et on appelle X et Y les numéros sortants.

1,5 **a)** Calculer l'entropie de leur somme $H(X+Y)$, ainsi que celle de leur différence $H(X-Y)$.

1,5 **b)** Montrer que l'application $(X, Y) \to (X + Y, X - Y)$ est bijective et en déduire la valeur de l'information mutuelle $I(X + Y, X - Y)$.

– EXERCICE 3. **Codage de source.**

Rappeler les définitions de code préfixe et de code uniquement décodable. Les codes ci-dessous sont-ils uniquement décodables ? Justifier.

1 **a)** $\{0, 10, 110, 111, 11111\}$,

1 **b)** $\{10, 11, 0101, 0000\}$,

1 **c)** $\{1, 110, 01, 010, 00000\}$.

– EXERCICE 4. **Longueur moyenne.**

Soit X une variable aléatoire prenant m valeurs avec une loi $p = (p_1 \dots, p_m)$. On code cette variable avec un code C préfixe de distribution des longueurs $(\ell_1, \dots, \ell_m)$. On suppose $\sum_{i=1}^{m} 2^{-\ell_i} = 1$ et on pose $q = (q_1, \dots, q_m)$ avec avec $q_i = 2^{-\ell_i}$.

0,5 **a)** Rappeler la définition de divergence de Kullback.

1,5 **b)** Montrer que la longueur moyenne du codage de X par le code C vaut

$$\overline{\ell} = H(p) + D(p \parallel q).$$

– EXERCICE 5. **Codage de Huffman.**

Soit X une variable aléatoire et C un codage de Huffman associé.

1,5 **a)** Rappeler une borne supérieure et une borne inférieure pour la longueur moyenne de C en fonction de $H(X)$. Justifier.

1,5 **b)** Soit $p = (0.09, 0.10, 0.11, 0.15, 0.25, 0.30)$ une loi sur X. Donner un codage de Huffman associé.