Master CSI 1 2015-2016

Théorie de l'information : DS du 19 octobre 2015

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

Responsable : Gilles Zémor

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

– Exercice 1. Soit u la loi uniforme $(1/m, 1/m, \ldots, 1/m)$ sur m objets. Quel est le maximum de la divergence de Kullback $\mathrm{D}(p \,||\, u)$, et pour quelle loi(s) p est-il atteint ?

– **Solution.**

$$\begin{aligned}\mathrm{D}(p \,||\, u) &= \sum_{i=1}^{m} p_i \log \frac{p_i}{1/m} \\ &= \sum_{i=1}^{m} p_i \log m + \sum_{i=1}^{m} p_i \log p_i \\ &= \log m - H(p).\end{aligned}$$

Le maximum de $\mathrm{D}(p \,||\, u)$ est donc atteint lorsque $H(p)$ est minimum, c'est-à-dire lorsque $H(p) = 0$, ce qui donne $\log m$ comme maximum de $\mathrm{D}(p \,||\, u)$ et qui arrive si et seulement si la loi p est une loi constante.

– Exercice 2. Pour f une fonction et X et Y deux variables aléatoires, expliquer pourquoi $H(X, Y, f(X, Y)) = H(X, Y)$ et en déduire que :

$$H(X, f(X, Y) \,|\, Y) = H(X \,|\, Y).$$

– **Solution.** On a $H(X, Y, f(X, Y)) = H(X, Y) + H(f(X, Y) \,|\, (X, Y))$. Or, comme $f(X, Y)$ est entièrement déterminé par (X, Y), on a $H(f(X, Y) \,|\, (X, Y)) = 0$. D'où

$$H(X, Y, f(X, Y)) = H(X, Y).$$

Par ailleurs,

$$\begin{aligned}H(X, Y, f(X, Y)) &= H(Y) + H(X, f(X, Y) \,|\, Y) \quad \text{et} \\ H(X, Y) &= H(Y) + H(X \,|\, Y).\end{aligned}$$

Comme on vient de montrer que les deux termes de gauche sont égaux, on en déduit

$$H(X, f(X,Y) \,|\, Y) = H(X \,|\, Y).$$

– Exercice 3. Soit $X = (X_1, X_2, X_3, X_4, X_5)$ une variable aléatoire prenant ses valeurs dans l'ensemble $\mathcal{X} \subset \{0,1\}^5$ à 10 éléments constitué de tous les quintuplets binaires de poids 2. On suppose la loi de $(X_1, X_2, X_3, X_4, X_5)$ uniforme.

a) Que vaut $H(X_i), i = 1..5$?

b) Que vaut $H(X_i, X_{i+1})$, où $i+1$ s'entend modulo 5 ?

c) Que vaut $H(X_{i+1} \,|\, X_i)$?

On crée maintenant la variable $Y = (Y_1, Y_2, Y_3, Y_4, Y_5)$ à partir de la variable X en remplaçant aléatoirement et uniformément un des trois bits 0 par un 1. Par exemple, on a :

$$\begin{aligned} P(Y = (11100) \,|\, X = (11000)) &= P(Y = (11010) \,|\, X = (11000)) \\ &= P(Y = (11001) \,|\, X = (11000)) \\ &= \frac{1}{3}. \end{aligned}$$

La variable Y prend donc ses valeurs dans l'ensemble $\mathcal{Y}$ des quintuplets binaires de poids 3.

Calculer $H(Y)$ et l'information mutuelle $I(X,Y)$.

– **Solution.**

a) Parmi les dix quintuplets de $\mathcal{X}$, quatre exactement ont un 1 en position i. Donc la loi de X_i est Bernoulli $(4/10, 6/10)$ et $H(X_i) = h(2/5)$.

b) En faisant défiler les dix quintuplets de $\mathcal{X}$, on voit en positions $i, i+1$ une fois le motif 11, trois fois le motif 10, trois fois le motif 01, et trois fois le motif 00. On a donc

$$H(X_i, X_{i+1}) = \frac{1}{10}\log_2 10 + 3\frac{3}{10}\log_2\frac{10}{3} = \log_2 10 - \frac{9}{10}\log_2 3.$$

c) On a $H(X_{i+1} \,|\, X_i) = H(X_i, X_{i+1}) - H(X_i)$, d'où

$$\begin{aligned} H(X_{i+1} \,|\, X_i) &= 1 + \log 5 - \frac{9}{10}\log 3 - \log 5 + \frac{2}{5} + \frac{3}{5}\log 3 \\ &= \frac{7}{5} - \frac{3}{10}\log 3. \end{aligned}$$

On a

$$\begin{aligned} P(Y = (11100)) &= \frac{1}{3}P(X = (11000)) + \frac{1}{3}P(X = (10100)) + \frac{1}{3}P(X = (01100)) \\ &= P(X = (11000)) = \frac{1}{10} \end{aligned}$$

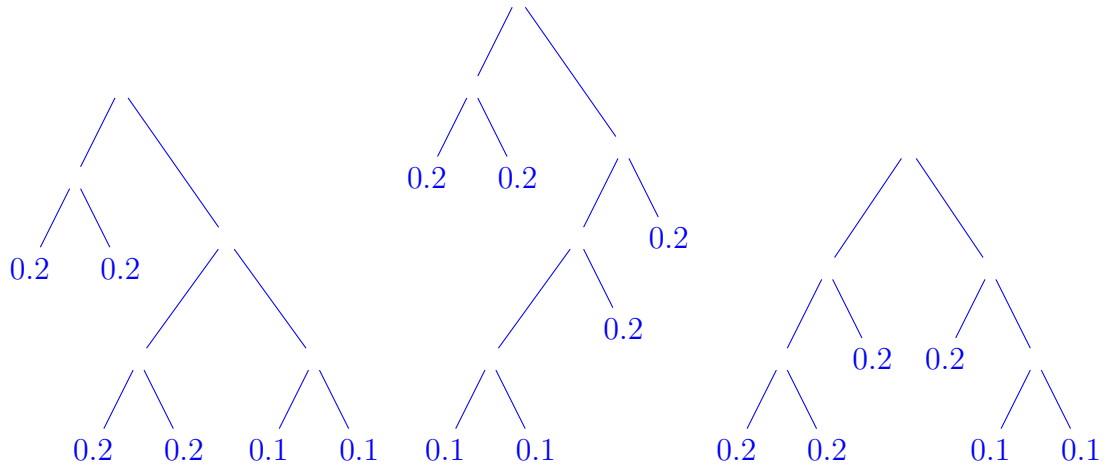
de même que pour les autres valeurs possibles de Y. La loi de Y est donc uniforme et $H(Y) = \log 10$.

On a par ailleurs $H(Y \mid X) = \log 3$, d'où

$$I(X, Y) = H(Y) - H(Y \mid X) = \log 10 - \log 3.$$

– Exercice 4. Soit la loi de probabilité $p = (0.2; 0.2; 0.2; 0.2; 0.1; 0.1)$. Donner tous les arbres de Huffman pour cette loi. Quelle est la longueur moyenne associée $\overline{\ell}$? Donner un code optimal pour cette même loi p qui n'est pas un code de Huffman.

– **Solution.** Les arbres donnés par l'algorithme de Huffman sont les suivants :

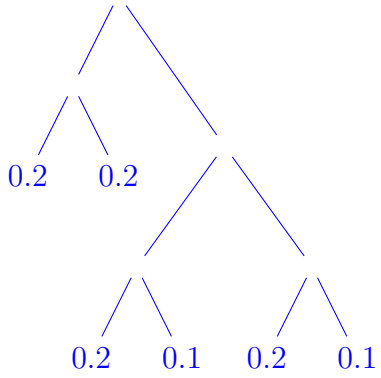


On obtient :

$$\overline{\ell} = 2.6.$$

Pour obtenir un code optimal qui n'est pas un code de Huffman, on peut échanger une des probabilités 0.1 avec une des probabilités 0.2 de même longueur. Ceci ne modifie pas la longueur moyenne, et on obtient que les deux probabilités les plus faibles ne sont pas issues d'un même sommet, contrairement à ce qu'exige l'algorithme de Huffman.

Ceci est possible sur le premier et le troisième arbre. Appliqué au premier arbre, on obtient par exemple, le code associé à l'arbre :



– Exercice 5. Étant donné une loi de probabilité p, on note $\overline{\ell}(p)$ la longueur moyenne d'un code de Huffman pour p. Soit P l'ensemble des lois de probabilité $p = (p_1, p_2, p_3, p_4)$ telles que :

$(1, 2, 3, 3)$ est la distribution des longueurs d'un code de Huffman pour p.

Que vaut le maximum λ de $\overline{\ell}(p)$ pour tous les $p \in P$? Quelles sont les lois $p \in P$ telles que $\overline{\ell}(p) = \lambda$?

– **Solution.** La loi P donne naissance à un code de Huffman de distribution des longueurs $(1, 2, 3, 3)$, avec la convention $p_1 \geqslant p_2 \geqslant p_3 \geqslant p_4$, si et seulement si $p_3 + p_4 \leqslant p_1$.

Il s'agit dans ces conditions de maximiser l'expression

$$\overline{\ell}(p) = p_1 + 2p_2 + 3(p_3 + p_4).$$

Ce maximum n'est clairement atteint que lorsque $p_3 + p_4 = p_1$ (sinon on peut augmenter $p_1 + 3(p_3 + p_4)$ sans modifier la somme $p_1 + p_3 + p_4$ et sans violer la condition $p_3 + p_4 \leqslant p_1$). On a alors

$$\overline{\ell}(p) = 4p_1 + 2p_2 = 4p_1 + 2(1 - p_1 - p_3 - p_4) = 4p_1 + 2(1 - 2p_1) = 2.$$

On a donc $\lambda = 2$ et, sachant que l'on doit respecter la contrainte

$$p_1 \geqslant p_2 \geqslant p_3 \geqslant \frac{1}{2}(p_3 + p_4) = p_1/2,$$

et que $p_2 = 1 - 2p_1$, on obtient que toutes les valeurs de p_1 acceptables sont dans l'intervalle $[\frac{1}{3}, \frac{2}{5}]$. Finalement, toutes les lois p acceptables sont les lois

$$p = (p_1, p_2, p_3, p_4) = (x, 1 - 2x, \frac{x}{2} + y, \frac{x}{2} - y)$$

avec $x \in [\frac{1}{3}, \frac{2}{5}]$ et $y \in [0, 1 - 5x/2]$.

– EXERCICE 6. Une urne contient b boules blanches et n boules noires. On réalise deux expériences : dans la première on tire une boule et on pose $X_1 = 0$ si la boule est blanche et $X_1 = 1$ si la boule est noire. Puis on remet la boule dans l'urne, et on recommence jusqu'à avoir formé le k-uple binaire aléatoire $X = (X_1, X_2, \ldots, X_k)$. Dans la deuxième expérience on tire k boules de l'urne, mais sans les remettre dans l'urne, de façon à créer un k-uple binaire $Y = (Y_1, Y_2, \ldots, Y_k)$.

Lequel des deux k-uples X ou Y a la plus grande entropie ? Justifier sans faire de calcul compliqué.

– **Solution.** Dans chacune des deux expériences, chacune des $b+n$ boules initialement dans l'urne à la même probabilité $1/(b+n)$ d'être sélectionnée au i-ième tirage. Donc

$$H(X_i) = H(Y_i) = h\left(\frac{b}{b+n}\right).$$

Les X_i sont clairement indépendantes, donc

$$H(X) = H(X_1) + \cdots + H(X_k).$$

Par contre, on peut écrire

$$\begin{aligned} H(Y) &= H(Y_1) + H(Y_2 \,|\, Y_1) + H(Y_3 \,|\, Y_1, Y_2) + \cdots H(Y_k \,|\, Y_1, Y_2, \ldots Y_{k-1}) \\ &< H(Y_1) + \cdots H(Y_k) = H(X) \end{aligned}$$

car la loi de Y_i dépend clairement de $Y_1, \ldots, Y_{i-1}$ et on doit avoir

$$H(Y_i \,|\, Y_1, \ldots Y_{i-1}) < H(Y_i).$$