Théorie de l'Information et Codage: Fiche d'exercices 2

à rendre pour le cours du 10 avril 2012.

Instructions: merci à chacun de rendre une copie manuscrite. Si vous avez réfléchi à plusieurs sur un problème, mettez les noms de vos collaborateurs.

Problème 1 (6 points): Codage Arithmétique

Une source sans mémoire émet des symboles dans l'alphabet $[0, K-1] = \{0, 1, \ldots K-1\}$ selon la loi: $P(U = i) = p(i) > 0$, pour tout $i \in [0, K-1]$. Le but du codage arithmétique est de compresser la source en binaire. On considère donc une suite $\mathbf{u} = (u_1, u_2, \ldots)$ dans $[0, K-1]^{\mathbb{N}}$ qui est compressée en $\mathbf{x} = (x_1, x_2, \ldots) \in \{0, 1\}^{\mathbb{N}}$. On notera $\mathbf{u}^{(m)} = (u_1, \ldots, u_m)$ le vecteur contenant les m premiers symboles de $\mathbf{u}$ (resp. $\mathbf{x}^{(m)}$ pour $\mathbf{x}$). De manière standard, on note $p(\mathbf{u}^{(m)}) = \prod_{i=1}^{m} p(u_i)$. On définit l'intervalle associé à $\mathbf{x}^{(n)} = (x_1, \ldots, x_n)$ par:

$$J(\mathbf{x}^{(n)}) = \left[\sum_{i=1}^{n} x_i 2^{-i}, \sum_{i=1}^{n} x_i 2^{-i} + 2^{-n} \right).$$

On définit $f_1(u) = \sum_{i=0}^{u-1} p(i)$ pour $u \leq K$ avec $f_1(0) = 0$, ainsi que pour $m \geq 1$,

$$f(\mathbf{u}^{(m)}) = f(\mathbf{u}^{(m-1)}) + f_1(u_m) p(\mathbf{u}^{(m-1)}) \text{ avec,}$$

$f(\mathbf{u}^{(0)}) = 0$ et $p(u^{(0)}) = 1$. On définit alors l'intervalle

$$J(\mathbf{u}^{(m)}) = \left[f(\mathbf{u}^{(m)}), f(\mathbf{u}^{(m)}) + p(\mathbf{u}^{(m)}) \right).$$

L'idée du codage arithmétique est d'encoder la source par $\mathbf{x}$ qui est l'écriture en binaire de la limite de $J(\mathbf{u}^{(m)})$ quand m tend vers l'infini: $\lim_{m \to \infty} J(\mathbf{u}^{(m)}) = \sum_{n=1}^{\infty} x_n 2^{-n}$.

1. Décrire un algorithme de décodage permettant de retrouver $\mathbf{u}$ à partir de $\mathbf{x}$ (on supposera que $\mathbf{x}$ ne se termine pas par une infinité de 0 ou de 1).

2. Quelle est la distribution de $\mathcal{X} = \sum_{n=1}^{\infty} x_n 2^{-n}$? Quelle est la distribution de $\mathcal{X}$ sachant $\mathbf{x}^{(n)}$? sachant $\mathbf{u}^{(m)}$?

Si $J(\mathbf{u}^{(m)})$ est inclus dans $J(\mathbf{x}^{(n)})$, alors l'encodeur peut émettre les n premiers bits de l'encodage de la source. Soit $M(n)$ le nombre minimum de lettres que la source doit émettre pour que l'encodeur puisse émettre les n premiers bits. $M(n)$ dépend de $\mathbf{u}$ et est donc une variable aléatoire.

3. Montrer que $p\left(\mathbf{u}^{(M(n))}\right) \leq 2^{-n}$ et que l'événement $\{M(n) \leq j\}$ ne dépend que de $\mathbf{u}^{(j)}$ et en particulier pas de $u_{j+1}, \ldots$. En déduire que

   $$n \leq E[M(n)] H(U)$$

   où $H(U)$ est l'entropie de la source.

4. Soit $p_{\min} = \min p(i)$, montrer que

$$E[M(n)]H(U) \le n + \log \frac{2e}{p_{\min}}.$$

Soit maintenant $N(m)$ le nombre minimum de bits émis par l'encodeur pour que le décodeur puisse décoder les m premiers symboles de la source.

5. Montrer que

$$-\log p(\mathbf{u}^{(m)}) \le E[N(m)|\mathbf{u}^{(m)}] \le -\log p(\mathbf{u}^{(m)}) + \log(4e),$$

et donc que $mH(U) \le E[N(m)] \le mH(U) + \log(4e)$.

6. Qu'en conclure sur les performances de l'encodage/décodage?

Problème 2 (12 points): Capacité zéro-erreur

On considère un canal discret sans mémoire de probabilité de transition $p(y|x)$ pour $(x, y) \in \mathcal{X} \times \mathcal{Y}$ et sans contrainte de coût, i.e. $b(x) = 0$. On note $C(p)$ sa capacité. Un code de longueur n est dit sans erreur si quelque soit le mot-code utilisé dans le canal la probabilité d'erreur est nulle, i.e. pour tout i, $P_E^{(i)} = 0$. Soit $M_0(n, p)$ la taille maximale d'un code de longeur n sans erreur. On définit la capacité zéro-erreur du canal par:

$$C_0(p) = \lim\sup_{n\to\infty} \frac{1}{n} \log M_0(n, p).$$

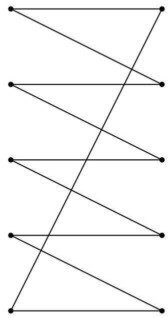
1. Montrer que $C_0(p) \le C(p)$ et que l'on a: $C_0(p) = \lim_{n\to\infty} \frac{1}{n} \log M_0(n, p) = \sup_n \frac{1}{n} \log M_0(n, p)$.

Soit G le graphe (d'adjacence) dont les sommets sont $\mathcal{X}$ et deux sommets x_1, x_2 sont connectés si il existe $y \in \mathcal{Y}$ tel que $p(y|x_1)p(y|x_2) > 0$. Pour simplifier, on prendra comme convention $\mathcal{X} = \{1, \ldots, |\mathcal{X}|\}$. On vérifiera que $C_0(p)$ ne dépend que de G. On notera dans la suite $C_0(G)$ à la place de $C_0(p)$.

2. Montrer que $C_0(G) = 0$ si et seulement si G est le graphe complet.

Un ensemble stable de G est un ensemble de sommets deux à deux non adjacents. On note $\alpha(G)$ la taille d'un stable maximum. Si G et H sont deux graphes, on définit leur produit $G \cdot H$ comme le graphe ayant pour sommets $V(G \cdot H) = V(G) \times V(H)$ et où (x_1, x_2) est adjacent à (y_1, y_2) si et seulement si pour $i = 1, 2$ $x_i = y_i$ ou x_i et y_i sont adjacents (dans G pour $i = 1$ et dans H pour $i = 2$). On note G^k le produit de k copies de G.

3. Montrer que $C_0(G) = \lim_{k\to\infty} \frac{1}{k} \log \alpha(G^k)$. Montrer que $\alpha(G^k) \ge \alpha(G)^k$. Montrer que pour le canal avec $|\mathcal{X}| = |\mathcal{Y}| = 5$ et dont les probabilités de transitions positives sont représentées par le graphe ci-dessous, l'inégalité précédente peut être stricte dès $k = 2$.



En déduire que $C_0(C_5) \geq \frac{1}{2}\log 5$, où C_5 est le pentagone, i.e. le cycle avec 5 sommets.

4. Montrer que si G peut être couvert par $\alpha(G)$ cliques alors $C_0(G) = \log \alpha(G)$. En déduire la valeur de $C_0(G)$ pour tous les graphes ayant au plus 5 sommets sauf pour le pentagone.

5. Soit $A_{ij} = \mathbf{1}(i = j \text{ ou } i \text{ et} j \text{ sont adjacents})$. Montrer que

$$-\log \min_{p_i} \sum_{i,j} A_{ij} p_i p_j \leq C_0,$$

où le minimum est pris sur les $p_i \geq 0$ avec $\sum_i p_i = 1$. On pourra introduire un (M, n)-code aléatoire bien choisi et montrer que la probabilité pour que les $M-1$ mots codes ne soient pas adjacents (dans G^n) à un mot code donné est

$$\left(1 - \left(\sum_{i,j} A_{ij} p_i p_j\right)^n\right)^{M-1} \geq 1 - M\left(\sum_{i,j} A_{ij} p_i p_j\right)^n,$$

puis on prendra $M = (1-\epsilon)^n \left(\sum_{i,j} A_{ij} p_i p_j\right)^{-n}$ et n suffisament grand pour conclure.

6. Montrer que la borne de la question précédente ne permet pas d'améliorer la borne inférieure déjà trouvée pour $C_0(C_5)$. Montrer que $C_0(C_5) \leq \log \frac{5}{2}$.

Nous allons maintenant montrer que $C_0(C_5) = \frac{1}{2}\log 5$, donc que la borne inférieure obtenue en 3. est la vraie valeur. Pour deux vecteurs $\mathbf{v} = (v_1, \ldots, v_n)^T$ et $\mathbf{w} = (w_1, \ldots, w_n)^T$, leur produit scalaire est noté $\mathbf{v}^T\mathbf{w} = \sum_i v_i w_i$. Pour deux vecteurs $\mathbf{v} = (v_1, \ldots, v_n)^T$ et $\mathbf{w} = (w_1, \ldots, w_m)^T$, leur produit tensoriel $\mathbf{v} \circ \mathbf{w}$ est le vecteur $(v_1w_1, \ldots v_1w_m, v_2w_1, \ldots v_nw_m)^T$ de longueur nm. On pourra vérifier que:

$$(\mathbf{x} \circ \mathbf{y})^T(\mathbf{v} \circ \mathbf{w}) = (\mathbf{x}^T\mathbf{v})(\mathbf{y}^T\mathbf{w}). \tag{1}$$

Une représentation orthonormale d'un graphe G ayant n sommets (notés $\{1, \ldots, n\}$) est un système $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ de vecteurs unitaires dans un espace Euclidien tels que si i et j ne sont pas adjacents alors $\mathbf{v}_i$ et $\mathbf{v}_j$ sont orthogonaux.

7. Soient $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ et $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ des représentations orthonormales de G et H respectivement. Montrer que les vecteurs $\mathbf{u}_i \circ \mathbf{v}_j$ forment une représentation orthonormale de $G \cdot H$.

On définit la valeur d'une représentation orthonormale $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ par

$$\min_{\mathbf{c}} \max_{1 \leq i \leq n} \frac{1}{(\mathbf{c}^T \mathbf{u}_i)^2},$$

où le minimum est pris sur les vecteurs $\mathbf{c}$ unitaires. On appelera le vecteur $\mathbf{c}$ atteignant le minimum le *manche*. On définit $\theta(G)$ la valeur minimale sur les représentations de G. On vérifiera que le minimum est atteint et on appelera la représentation correspondante optimale.

8. Montrer que $\theta(G \cdot H) \leq \theta(G)\theta(H)$.
9. Montrer que $\alpha(G) \leq \theta(G)$.
10. Montrer que $C_0 \leq \log \theta(G)$. En déduire $C_o(C_5)$.

Problème 3 (2 points):
Etant donné deux canaux sans mémoire ayant des probabilités de transition $p' : \mathcal{X}' \to \mathcal{Y}'$ et $p'' : \mathcal{X}'' \to \mathcal{Y}''$, on définit le canal 'produit' qui a pour probabilités de transition: $p : \mathcal{X}' \times \mathcal{X}'' \to \mathcal{Y}' \times \mathcal{X}''$ définies par

$$p(y', y''|x', x'') = p'(y'|x')p''(y''|x'').$$

1. Donner en quelques mots un interprétation du canal produit. Calculer la capacité du canal produit en fonction des capacités C' et C'' des deux canaux initiaux.
2. Montrer que si on considère les capacités zéro-erreur de ces canaux, on a $C_0 \geq C_0' + C_0''$. Pour certains canaux, l'inégalité peut être stricte (il N'est PAS demandé de donner un exemple). Cependant si le graphe d'adjacence (défini dans l'exercice précédent) du premier canal G' peut être couvert par $\alpha(G')$ cliques alors montrer que l'inégalité ne peut pas être stricte.