

# Solutions des Exercices du cours de Théorie de l'Information et Codage

## cours 1 du 17 février 2009.

1. Dans le cas  $R = 1/2n$ , avec  $n \in \mathbb{N}$  et pour un encodage du type: répétition de chaque bit  $2n$  fois sur le canal BSC: donner une stratégie de décodage et calculer la probabilité d'erreur correspondante  $P_e$ . Comparer au cas  $R = 1/(2n - 1)$  étudié en cours quand  $p \rightarrow 0$ .

- Décodage par majorité, en cas d'égalité, tirage à pile ou face.

$$P_e = \sum_{k \geq n+1} \binom{2n}{k} p^k (1-p)^{2n-k} + \frac{1}{2} \binom{2n}{n} p^n (1-p)^n.$$

- quand  $p \rightarrow 0$ , on a  $P_e \sim \binom{2n-1}{n} p^n$  qui correspond au cas étudié en cours.

2. On considère le cas  $R = 2n + 1$  pour  $n \in \mathbb{N}$  et l'encodage: j'envoie la majorité de chacun des blocs de  $2n + 1$  bits successifs émis par la source (comme décrit en cours pour  $R = 3$ ). Montrer que  $P_e = (1 - p)Q + p(1 - Q)$  avec

$$Q = \frac{1}{2} - \binom{2n}{n} 2^{-(2n+1)}.$$

Montrer que dans le cas  $R = 2n$ , une stratégie similaire donne exactement la même probabilité d'erreur  $P_e$ .

- Soit  $X_1, X_2, \dots$  une suite de v.a. indépendantes de Bernoulli de paramètre  $1/2$ . Si  $\sum_{i=1}^{2n+1} X_i \leq n$ , j'envoie 0 et si la transmission sur le canal se fait sans erreur, le décodeur fait  $\sum_{i=1}^{2n+1} X_i$  erreurs, soit un nombre d'erreurs moyen de:

$$\mathbb{E} \left[ \sum_{i=1}^{2n+1} X_i; \sum_{i=1}^{2n+1} X_i \leq n \right] = \left( \frac{1}{2} \right)^{2n+1} \sum_{k=1}^n k \binom{2n+1}{k}.$$

On a  $\sum_{k=1}^{2n+1} k \binom{2n+1}{k} = (2n+1)2^{2n} = 2 \sum_{k=1}^n k \binom{2n+1}{k} + (n+1) \binom{2n+1}{n+1}$  et  $(n+1) \binom{2n+1}{n+1} = (2n+1) \binom{2n}{n}$ . On a donc

$$\begin{aligned} \mathbb{E} \left[ \sum_{i=1}^{2n+1} X_i; \sum_{i=1}^{2n+1} X_i \leq n \right] &= \left( \frac{1}{2} \right)^{2n+1} (2n+1) \left( 2^{2n} - \binom{2n}{n} \right) \\ &= (2n+1) \frac{Q}{2}. \end{aligned}$$

Par symétrie, le nombre d'erreurs par bits est

$$\frac{2}{2n+1} \mathbb{E} \left[ \sum_{i=1}^{2n+1} X_i; \sum_{i=1}^{2n+1} X_i \leq n \right] = Q.$$

Au final, si la transmission sur le canal est correcte (avec probabilité  $1 - p$ ), le nombre moyen d'erreurs par bit est  $Q$  et sinon (probabilité  $p$ ), le nombre moyen d'erreurs par bit est  $1 - Q$ , d'où  $P_e = (1 - p)Q + p(1 - Q)$ .

- pour  $R = 2n$ , en appliquant la stratégie: majorité sur des blocs de longueur  $2n$ . Si  $\sum_{i=1}^{2n} X_i = n$ , dans tous les cas, le nombre d'erreurs est  $n$ . On a alors avec le même raisonnement:

$$\begin{aligned} \mathbb{E} \left[ \sum_{i=1}^{2n} X_i; \sum_{i=1}^{2n} X_i \leq n-1 \right] &= \left(\frac{1}{2}\right)^{2n} \sum_{k=1}^{n-1} k \binom{2n}{k} \\ &= \left(\frac{1}{2}\right)^{2n+1} \left( 2n2^{2n-1} - n \binom{2n}{n} \right) \\ &= nQ. \end{aligned}$$

Le nombre d'erreur par bits est alors:

$$\frac{2}{2n} \mathbb{E} \left[ \sum_{i=1}^{2n} X_i; \sum_{i=1}^{2n} X_i \leq n-1 \right] + \left(\frac{1}{2}\right)^{2n+1} \binom{2n}{n} = Q + \left(\frac{1}{2}\right)^{2n+1} \binom{2n}{n}.$$

Donc au final on obtient une probabilité d'erreur  $(1-p)Q + p(1-Q) + \left(\frac{2n}{n}\right)2^{-(2n+1)} > P_e!$   
Il vaut donc mieux envoyer la majorité de blocs de longueur  $2n+1$  comme précédemment (en envoyant moins de symboles dans le canal).

3. Vérifier les formules pour les probabilités d'erreur par bit données dans le cours pour le code de Hamming.

- On a vu que si une seule erreur est faite sur les 7 bits, elle est corrigée. On considère maintenant le premier digit  $x_1$  et on calcule la probabilité de faire une erreur  $p_1 = \mathbb{P}(\hat{x}_1 \neq x_1)$  en fonction du nombre d'erreurs total. Si 7 erreurs sont faites avec probabilité  $p^7$  alors le syndrome est  $s = (0, 0, 0)$  et on a bien  $\hat{x}_1 \neq x_1$ . On considère le cas de 6 erreurs.

$$H \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} + H \begin{pmatrix} 0 \\ 1 \\ \dots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \end{pmatrix} \text{ on a: } H \begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix} = H \begin{pmatrix} 0 \\ 1 \\ \dots \\ 1 \end{pmatrix}.$$

Donc si 6 erreurs sont faites, le syndrome correspond à la 7-ième colonne de  $H$  et donc le décodeur change le seul digit qui avait bien été envoyé et donc rajoute une erreur. Donc pour les 7 motifs d'erreurs de poids 6, on aura  $\hat{x}_1 \neq x_1$ , ce qui correspond à un événement de probabilité  $7p^6(1-p)$ . Dans le cas où le motif d'erreur a poids 5, un raisonnement similaire montre que le décodeur corrige une erreur parmi les 5. Donc pour que  $\hat{x}_1 \neq x_1$ , il faut qu'il y ait eu une erreur sur le premier digit (soit  $\binom{6}{4}$  possibilités) et qu'elle ne soit pas corrigée (3 possibilités) soit un événement de probabilité  $12p^5(1-p)^2$ . On considère maintenant que le motif d'erreur a poids 4. Si les 3 colonnes complémentaires du motif d'erreur sont liées, alors le syndrome est  $(0, 0, 0)$  et les erreurs ne sont pas corrigées. Il y a 3 triplets de colonnes liées qui contiennent une colonne donnée et 4 triplets qui ne contiennent pas une colonne donnée. Dans ce cas pour que  $\hat{x}_1 \neq x_1$ , il faut qu'il y ait une erreur sur le premier digit et que les 3 colonnes complémentaires soient liées soit une probabilité  $4p^4(1-p)^3$ . Si les 3 colonnes complémentaires du motif d'erreur sont libres, alors une des erreurs est corrigée. Dans ce cas, pour que  $\hat{x}_1 \neq x_1$ , il faut donc qu'il y ait une

erreur sur le premier digit et que les colonnes soient libres ( $\binom{3}{6} - 4 = 16$  possibilités) et que l'erreur ne soit pas corrigée ( $16 - 4 = 12$  possibilités), soit une probabilité de  $12p^4(1-p)^3$ . En sommant les contributions dans le cas de 4 erreurs on a bien  $16p^4(1-p)^3$ . Le cas de 3 erreurs est similaire. Cette fois si les 3 colonnes correspondant au motif d'erreur sont liées, le syndrome est  $(0, 0, 0)$  et si elles sont libres, le décodeur va rajouter une erreur. Soit une contribution de  $(3 + 12 + 4)p^3(1-p)^4$ . Dans le cas de 2 erreurs, le décodeur va rajouter une erreur, ce qui donne une contribution de  $(6 + 3)p^2(1-p)^5$ .

4. Soit  $X$  une v.a. discrète à espace d'états  $E$  et  $f$  une fonction réelle définie sur  $E$ . Montrer que  $H(X) \geq H(f(X))$ , avec égalité ssi  $f$  restreinte à  $\{x : \mathbb{P}(X = x) > 0\}$  est une injection.

• On a

$$\begin{aligned} H(X) - H(f(X)) &= - \sum_x p(x) \log p(x) + \sum_y \left( \sum_{x, f(x)=y} p(x) \right) \log \sum_{x, f(x)=y} p(x) \\ &= - \sum_x p(x) \log \frac{p(x)}{\sum_{x', f(x')=f(x)} p(x')} \geq 0, \end{aligned}$$

avec égalité si pour  $x$  tel que  $p(x) > 0$ ,  $p(x) = \sum_{x', f(x')=f(x)} p(x')$ , c'est à dire  $f(x') = f(x)$  implique  $x = x'$ .

5. Etant donnés une probabilité  $(p_1, p_2, \dots, p_n)$  et un entier  $m \leq n$ , on définit  $q_m = 1 - \sum_{j=1}^m p_j$ . Montrer que  $H(p_1, \dots, p_n) \leq H(p_1, \dots, p_m, q_m) + q_m \log(n - m)$ . Cas d'égalité?

• le cas  $n = m$  est trivial avec  $0 \log 0 = 0$ . Pour  $m < n$ , on a:

$$\begin{aligned} H(p_1, \dots, p_n) - H(p_1, \dots, p_m, q_m) &= \sum_{i=m+1}^n p_i \log \frac{\sum_{j=m+1}^n p_j}{p_i} \\ &= q_m \sum_{i=m+1}^n \frac{p_i}{\sum_{j=m+1}^n p_j} \log \frac{\sum_{j=m+1}^n p_j}{p_i} \\ &\leq q_m \log(n - m), \end{aligned}$$

avec égalité ssi  $p_i = \frac{q_m}{n-m}$ .

6. Soit  $f(x)$  une fonction définie pour tout  $x \geq 1$ . Si  $X$  est une v.a. discrète à espace d'états  $E = \{x_1, \dots, x_n\}$ , on définit la  $f$ -entropie de  $X$  par  $H_f(X) = \sum_{i=1}^n p_i f(1/p_i)$ , où  $p_i = \mathbb{P}(X = x_i)$ . Si  $f$  est concave, trouver la meilleure borne supérieure pour  $H_f(X)$  qui ne dépende que de  $n$ . Si  $f(x) = \log(x)/x$ , montrer que  $H_f(X) < \log(e)/e$ . Montrer qu'en fait,  $H_f(X) \leq \log(3)/3$ , avec égalité ssi exactement trois  $p_i$  sont égaux à  $1/3$  et le reste à 0.

• Par l'inégalité de Jensen, on a  $H_f(X) \leq f(n)$ .

• Soit  $m$  le nombre de  $p_i$  non nuls. On a

$$\begin{aligned} H_f(X) &= \sum_i p_i^2 \log p_i^{-1} \\ &\leq \left( \sum_i p_i^2 \right) \log \frac{1}{\sum_i p_i^2}, \end{aligned}$$

avec égalité ssi  $p_i^2 / \sum p_i^2 = 1/m$ . La fonction  $x \mapsto -x \log x$  atteint un maximum en  $x = e^{-1}$  sur  $\mathbb{R}_+$  égal à  $\log(e)/e$ . Pour le cas d'égalité, on a  $p_i = 1/m$  pour  $i \leq m$  et  $p_i = 0$  pour  $i > m$  et dans ce cas  $H_f(X) = \frac{1}{m} \log m \leq \log(3)/3$ . Il faut montrer que  $\max \sum_i -p_i^2 \log p_i \leq \log(3)/3$  sous les contraintes  $\sum_i p_i = 1$  et  $p_i \geq 0$ . Soit  $p^*$  un point où le maximum est atteint. Si il existe  $0 < p_j^* < p_i^*$ , alors on a au premier ordre en  $\epsilon$ :

$$H_f(p_i^* - \epsilon, p_j^* + \epsilon) - H_f(p^*) = 2\epsilon(\log p_i^* - \log p_j^*) > 0,$$

ce qui est contradictoire. On a donc  $p_i^* = 1/m$  pour  $i \leq m$  et  $p_i^* = 0$  pour  $i > m$ , cas étudié auparavant.

7. Dans l'exemple 4 du cours (DMC  $Y = X^2$ ), montrer que  $Cov(X, Y) = 0$  et  $I(X; Y) = 1$  bit.

- Rappel  $Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$  donc ici  $Cov(X, Y) = \mathbb{E}[XY] = 0$ . On a  $I(X; Y) = 4\frac{1}{4} \log \frac{1/2}{1/4} = 1$ .

8. Trouver des conditions nécessaires et suffisantes sur la distribution jointe  $p(x, y, z)$  pour qu'il y ait égalité dans le Théorème 1.2. Pour tout  $r \geq 2$  et toute valeur  $P_e \in [0, 1]$ , construire des v.a.  $X$  et  $Y$  à valeurs dans  $\{1, \dots, r\}$  pour lesquelles l'inégalité de Fano est optimale.

- CNS: pour tout  $z$  tel que  $p(x, y, z) > 0$ , la loi du couple  $(X, Y)$  sachant  $Z = z$  doit être uniforme. Donc la construction suivante donne une égalité dans l'inégalité de Fano: tirer  $Z$  selon une Bernoulli de paramètre  $P_e$ . Si  $Z = 0$ , prendre  $X = Y = x_i$  avec probabilité  $1/r$ . Si  $Z = 1$ , prendre pour loi de  $(X, Y)$  la loi uniforme sur  $\{x_1, \dots, x_r\}^2 \setminus \{(x_1, x_1), \dots, (x_r, x_r)\}$ .

9. Montrer que l'inégalité de Fano donne une borne supérieure et une borne inférieure pour  $P_e$  en fonction de  $H(X|Y)$ . Donner une interprétation heuristique pour la borne supérieure.

- Soit  $f(x) = -x \log x - (1-x) \log(1-x) + x \log(r-1)$ . L'inégalité de Fano donne:

$$f(P_e) \geq H(X|Y). \tag{1}$$

On a  $f'(x) = \log(1-x) - \log x + \log(r-1)$ , soit  $f'(x) < 0 \Leftrightarrow (r-1)/r < x$ . Donc le maximum de  $f$  est  $f((r-1)/r) = \log(r)$ , comme  $H(X|Y) \leq H(X) \leq \log r$ , l'inégalité (1) donne bien une borne inférieure et supérieure sur  $P_e$ . La borne supérieure est non triviale uniquement lorsque  $H(X|Y) > \log(r-1)$ . En effet dans ce cas, on doit avoir  $P_e < 1$  puisque si  $X \neq Y$ , la quantité d'information nécessaire pour déterminer  $X$  parmi les  $r-1$  valeurs restantes est au plus  $\log(r-1)$ .

10. Si  $X, Y$  et  $Z$  sont des v.a. discrètes, on définit l'information mutuelle conditionnelle  $I(X; Y|Z)$  comme suit

$$I(X; Y|Z) = \sum_{x,y,z} p(x, y, z) \log \frac{p(x, y|z)}{p(x|z)p(y|z)}.$$

Montrer que  $I(X; Y|Z) = I(Y; X|Z)$ ;

$I(X; Y, Z) = I(X; Z) + I(X; Y|Z)$ ;

$I(X; Y|Z) \geq 0$  avec égalité ssi  $(X, Z, Y)$  est une chaîne de Markov.

- Le premier point est clair. Comme vous l'aurez remarqué, le deuxième point était faux dans l'énoncé initial...

$$\begin{aligned}
 I(X; Z) + I(X; Y|Z) &= \sum p(x, y, z) \log \frac{p(z|x)}{p(z)} \frac{p(x, y|z)}{p(x|z)p(y|z)} \\
 &= \sum p(x, y, z) \log \frac{p(x, y, z)}{p(x)p(yz)} = I(X; Y, Z).
 \end{aligned}$$

Pour le dernier point:

$$\begin{aligned}
 -I(X; Y|Z) &= \sum p(x, y, z) \log \frac{p(x|z)p(y|z)}{p(x, y|z)} \\
 &\leq \log \sum \frac{p(x, z)p(y, z)}{p(z)} \\
 &= \log \sum_{y, z} \frac{p(z)p(y, z)}{p(z)} = 0,
 \end{aligned}$$

avec égalité ssi  $p(x, y|z) = p(x|z)p(y|z)$ , comme  $p(x, y|z) = p(x, y, z)/p(z) = p(x|y, z)p(y|z)$ , cela signifie que  $p(x|z) = p(x|y, z)$ , c'est à dire que  $(X, Z, Y)$  est une chaîne de Markov.