Exercices du cours de Théorie de l'Information et Codage
cours 3 du 3 mars 2009.

1. Soient $X_1, \ldots, X_n$, des v.a. discrètes. Montrer que

$$H(X_1, \ldots, X_n) = H(X_1) + \sum_{k=2}^{n} H(X_k|X_1, \ldots, X_{k-1}).$$

2. On considère une chaîne de Markov (à espace d'états finis) ergodique $S_0, S_1, \ldots$ avec un état initial arbitraire. Montrer que $H(S_2|S_1, S_0) = H(S_2|S_1)$ et que pour tout $n \geq 2$

$$H(S_1, S_2, \ldots S_n|S_0) = \sum_{k=1}^{n} H(S_k|S_{k-1}).$$

Simplifier la formule quand S_0 suit la loi stationnaire. Pour une source Markovienne $X_1, X_2, \ldots$ montrer que $H(X_1, \ldots X_n|S_0) = H(S_1, \ldots S_n|S_0)$. Montrer que pour une source Markovienne stationnaire:

$$H(X_1, \ldots X_n|S_0) = nH(X|S).$$

Interpréter le résultat précédent pour le codage par blocs d'une source Markovienne.

3. On considère le problème d'encodage d'une suite binaire $x^n \in \{0, 1\}^n$.

   - Algorithme 1: nous lisons entièrement la suite, puis nous procédons en deux étapes: nous envoyons d'abord le nombre de 1 dans la suite puis nous envoyons les emplacements de ces 1 dans la suite. Montrer que la longueur du mot code est bornée par

   $$\ell(x^n) \leq H(k/n) + 1/2 \log n - 1/2 \log \left(\pi \frac{k(n-k)}{n^2}\right) + 3,$$

   où $k = \sum_i x_i$. On pourra utiliser l'encadrement suivant:

   $$\sqrt{\frac{n}{8k(n-k)}} \leq \binom{n}{k} 2^{-nH(k/n)} \leq \sqrt{\frac{n}{\pi k(n-k)}}.$$

   - Algorithme 2: utiliser un code arithmetique avec la règle de Laplace pour le modèle prédictif. Montrer que la longueur du mot code est bornée par

   $$\ell(x^n) \leq H(k/n) + 1/2 \log n - 1/2 \log \left(\pi \frac{k(n-k)}{n^2}\right) + 2.$$

Quel est l'intérêt de la seconde méthode?