L3 MIA / Systèmes de Communications

TD Codage de source

1 Codage de Huffman

Une source sans mémoire X émet 4 symboles x_1, x_2, x_3 et x_4 avec les probabilité suivantes : $\mathrm{P}(x_1) = 1/2$, $\mathrm{P}(x_2) = 1/4$ et $\mathrm{P}(x_3) = \mathrm{P}(x_4) = 1/8$.

a) Construire un code de Huffman pour X. Montrer que chaque symbole porte un bit d'information par élément binaire, *i.e.* $I(x_i) = n_i$, où n_i est le nombre d'éléménts binaires du symbole x_i.

b) Calculer l'entropie de la source, la longueur moyenne des mots de code et l'efficacité de ce codage.

c) Comparer l'efficacité de ce code à celle d'un code de longueur fixe (2 éléments binaires par symboles). Combien d'information porte chaque élément binaire ?

2 L'union fait l'efficacité

Une source sans mémoire ne prend que deux valeurs x_0 et x_1, de probabilités respectives $\mathrm{P}(x_0) = 0.7$ et $\mathrm{P}(x_1) = 0.3$.

a) Calculer l'entropie de cette source. En déduire la quantité d'information par élément binaire si l'on code x_0 par 0 et x_1 par 1.

b) On décide à présent de coder les symboles par paquets de 2. Calculer les probabilités de chacune des 4 combinaisons, construire un codage de Huffman. Combien de bits d'information porte alors chaque élément binaire ?

De manière générale, si l'on regroupe les symboles d'une source par paquets de N, alors il existe un code tel que la longueur moyenne par symbole L/N vérifie :

$$H(X) \leq \frac{L}{N} \leq H(X) + \frac{1}{N}$$

où $H(X)$ est l'entropie par symbole. Ainsi, grouper les symboles avant de les coder permet de s'approcher de l'efficacité maximale.