Introduction à la théorie de l'information

TD 6 : Introduction au codage source

Université Paris 13, Institut Galilée
Master Ingénierie et Innovations en Images et Réseaux - 1ère année

2017-2018

Exercice 1 : Codes non singuliers, déchiffrables et instantanés

On considère une source sans mémoire X à valeurs dans $\mathcal{X} = \{x_1, x_2, x_3, x_4\}$ et les codes suivants :

Symbole	Code A	Code B	Code C	Code D	Code E	Code F
x_1	0	0	00	0	0	0
x_2	11	11	01	10	01	11
x_3	00	00	10	110	011	101
x_4	11	01	11	1110	0111	100

1. Dire pour chacun de ces codes, s'ils sont non singuliers, déchiffrables et/ou instantanés.
2. Vérifier que les codes instantanés vérifient bien l'inégalité de Kraft
3. Proposer un code G instantané pour la source, différent de ceux présentés ci-dessous

Exercice 2 : Construction d'un code instantané

On considère une source sans mémoire associée à une variable aléatoire $X \in \{A, B, C, D, E\}$. On souhaite la coder grâce à un code binaire.

1. Pourra-t-on la coder avec un code instantané ayant 2 mots de 4 bits, 1 mot de 3 bits, 1 mot de 2 bits et 1 mot de 1 bit ?
2. Si oui, construire un tel code.
3. Si l'on a les probabilités d'apparitions suivantes, quelle est la correspondance symboles/mots que l'on doit utiliser pour minimiser la longueur moyenne du code ?

A	**B**	**C**	**D**	**E**
0.1	0.5	0.2	0.1	0.1

Exercice 3 : Codage de Shannon-Fano, codage de Huffman

Soit X une source à valeurs dans $\{A, E, G, I, L\}$ où les probabilités d'apparition des symboles sont :

A	**E**	**G**	**I**	**L**
$\frac{12}{31}$	$\frac{5}{31}$	$\frac{4}{31}$	$\frac{6}{31}$	$\frac{4}{31}$

1. Construire le code de Shannon-Fano associé à la variable aléatoire X et calculer la longueur moyenne du code.
2. Même question avec un code de Huffman.
3. Lequel réalise la meilleure compression ?

Exercice 4 : Intérêt du code de Huffman

On considère une source sans mémoire quaternaire associée à une variable aléatoire $X \in \{-3, -1, 1, 3\}$. On souhaite associer à chaque symbole un message binaire. Nous allons comparer ici un codage binaire de Gray ne tenant pas compte des caractéristiques de la source, et un code de Huffman prenant en compte les spécificités de la source.

1. On suppose que l'on connaît une réalisation typique de la source :

$$y = [-3, 1, 3, 3, -1, 1, 3, 3, -3, 1, 3, -1, 3, 1, -1, 3]$$

   Calculer la loi de probabilité $p_X(x)$ de X et calculer l'entropie $H(X)$ de la source.

2. On considère dans un premier temps un code binaire de Gray : $-3 \to 00, -1 \to 01, 1 \to 11, 3 \to 10$. Calculer la longueur moyenne, le rendement et la redondance du code. Coder le message y et donner sa taille (en bits).

3. Construire le code de Huffman associé à la source. Calculer la longueur moyenne, le rendement et la redondance du code. Coder le message y et donner sa taille (en bits).

4. Conclure.

Exercice 5 : Codage d'un mot

Coder grâce à un codage d'Huffman le mot INSTITUT_GALILEE et calculer le rendement et la redondance de ce code.

Exercice 6 : Codage de l'extension d'une source

On considère une source sans mémoire X à valeurs dans $\{A, B\}$. On suppose que $p_X(A) = \frac{1}{4}$.

1. On considère la source X.
   (a) Calculer son entropie $H(X)$
   (b) Réaliser un codage de la source X grâce à un code de Huffman
   (c) Calculer la longueur moyenne, le rendement et la redondance du code
   (d) Quel est le nombre moyen de bits nécessaire pour coder un symbole ?

2. Même question pour l'extension de degré 2 de X, notée $X^{[2]}$

3. Même question pour l'extension de degré 3 de X, notée $X^{[3]}$

4. Quelle est la taille de bloc minimale que l'on doit utiliser pour être sur que la longueur moyenne du code pour coder un symbole soit inférieure à 0.82 ?