

Codage de l'information

Examen de juin 2005 (2nde session)

durée 1h30 - documents et calculatrices non autorisés

Exercice 1 : *Code* (1/2H)

On considère le langage L défini sur l'alphabet $\mathcal{A} = \{0, 1\}$, n et m étant deux entiers naturels, par

$$L = \{0, 1^n 0, 1^m\}$$

où la notation 1^n désigne le mot constitué de n fois la lettre 1.

Q 1 . À quelle condition portant sur les entiers n et m le langage L est-il un code ?

Q 2 . Dans cette question, on pose $n = 3$ et $m = 2$. On note A le sous-ensemble de \mathbb{N} dont l'écriture en base 2 est un mot binaire qui se décompose en mots de L .

Q 2.1. Parmi les entiers de 0 à 10, quels sont ceux qui appartiennent à A ?

Q 2.2. Y a-t-il des puissances de 2 dans A ?

Q 2.3. Y a-t-il des nombres de la forme $2^k - 1$ dans A ? si oui pour quels entiers k ?

Exercice 2 : *Fichier à compresser* (1/2H)

Un fichier sur disque est constitué de 1500 octets. On s'aperçoit qu'il contient beaucoup plus de 1 que de 0 et on souhaite le compresser. Pour cela on le décompose en 3000 mots de 4 bits et on constate qu'il y a alors :

- $x = 1200$ mots de la forme 1111,
- $y = 300$ mots pour chacun des quatre motifs ne contenant qu'un seul 0 (0111, 1011, 1101, 1110),
- $z = 100$ mots pour chacun des six motifs contenant deux 0 (0011, 0101, 0110, 1001, 1010, 1100),
- aucun mot contenant plus de deux 0.

Q 1 . Comment coder ces 11 motifs pour obtenir une compression maximale ? Quelle est alors la taille en bits du fichier ainsi comprimé ?

Q 2 . Sans la calculer explicitement, donner un encadrement de l'entropie de ce fichier considéré comme source de symboles de quatre bits.

Exercice 3 : *Détection et correction d'erreurs* (1/2H)

On considère le codage suivant des mots de 4 bits en des mots de huit bits :

$$\begin{array}{ccc} \mathbf{c} : & \mathbb{F}_2^4 & \longrightarrow \mathbb{F}_2^8 \\ \mathbf{u} = \langle b_1, b_2, b_3, b_4 \rangle & \longmapsto & \mathbf{c}(\mathbf{u}) = \langle b_1, b_2, b_3, b_4, b_1 \oplus b_2, b_3 \oplus b_4, b_1 \oplus b_3, b_2 \oplus b_4 \rangle \end{array}$$

Q 1 . Calculez $\mathbf{c}(\langle 0, 0, 0, 0 \rangle)$, $\mathbf{c}(\langle 0, 1, 0, 0 \rangle)$, $\mathbf{c}(\langle 1, 0, 1, 1 \rangle)$.

Q 2 . Vérifiez que \mathbf{c} est bien un codage, i.e. est une application injective.

Q 3 . On désigne par C le sous-ensemble de \mathbb{F}_2^8 constitué des 16 mots images par \mathbf{c} d'un mot de \mathbb{F}_2^4 . Montrez que C est un code linéaire.

Q 4 . Quelle est la distance minimale de ce code ? En déduire ces capacités de détection et de correction d'erreurs.