

Codage de l'information**Devoir surveillé n° 1**

Lundi 5 novembre 2012 - Durée 2h - Documents non autorisés. Calculatrices autorisées.

Veuillez indiquer le numéro de votre groupe de TD sur la copie qu'il est inutile de rendre anonyme, ainsi que votre NIP (figurant sur votre carte d'étudiant).

Ce sujet contient trois exercices indépendants. Prenez 10mn pour lire l'intégralité du sujet avant de commencer.

Exercice 1-1 *Représentation binaire des entiers*

On note N le nombre entier naturel dont l'écriture décimale comprend quatre chiffres, le plus à gauche étant un 3 et les trois suivants étant les trois derniers chiffres de votre NIP.

$$N = 3xyz$$

Question 1 Sans les calculer, donnez la taille des représentations binaires et hexadécimales de N .

Question 2 Donnez les écritures binaires et hexadécimales de N .

Question 3 Dans le cadre d'une représentation signée en complément à deux sur 12 bits, quel est l'entier représenté par l'écriture binaire de N ?

Question 4 Se peut-il que l'un des étudiants composant ce matin, trouve (avec raison) un entier d'un signe différent à celui que vous avez trouvé dans la question précédente ?

Exercice 1-2 *Opérations logiques sur les entiers*

On considère la fonction f définie sur les entiers naturels par

$$f(0) = 0,$$

et pour tout entier $n > 0$

$$f(n) = n \wedge (n - 1).$$

Question 1 Calculez $f(N)$ où N est l'entier de l'exercice précédent.

Question 2 Que vaut $f(n)$ si n est une puissance de 2 ?

Question 3 Comparez les écritures binaires de n et $f(n)$ lorsque n est un entier non nul. En particulier que devient le bit 1 le plus à droite dans n ?

Question 4 Que calcule l'algorithme 1.1 (cf page 2) ?

Exercice 1-3 *Code génétique*

Le code génétique est le code utilisé dans l'ADN, au cœur de nos cellules. Il s'agit d'un code à longueur fixe. Il existe un codage qui associe un mot du code génétique à un acide aminé. On rappelle que l'alphabet de l'ADN est constitué de quatre lettres (A, C, G et T) et qu'il existe 21 acides aminés. Il existe donc (au moins) 21 mots dans le code génétique.

Question 1 Quelle doit être la longueur minimale des mots du code génétique ?

Algorithme 1.1 Algorithme de la question 4

Entrée : $n \in \mathbb{N}$ **Sortie :** ??? $p \leftarrow 0$ **tant que** $n \neq 0$ **faire** $p \leftarrow p + 1$ $n \leftarrow f(n)$ **fin tant que****renvoyer** p

Question 2 Pourrait-on trouver un autre code, dont 1 mot serait de longueur 1, 10 de longueurs 2 et 10 de longueurs 3? Et un code préfixe?

Question 3 Mêmes questions avec 14 mots de longueur 2 et 7 mots de longueur 3?

Exercice 1-4 *Miroirs et suffixes*

Dans cet exercice, \mathcal{A} désigne un alphabet.

Soit $\mathbf{u} \in \mathcal{A}^*$ un mot. Le *miroir* du mot \mathbf{u} est un mot, noté $\tilde{\mathbf{u}}$ dont les lettres parcourues de gauche à droite sont celles de \mathbf{u} parcourues de droite à gauche. Ainsi, si

$$\mathbf{u} = x_1 x_2 \dots x_n,$$

où les x_i sont des lettres, alors

$$\tilde{\mathbf{u}} = x_n \dots x_2 x_1.$$

Le miroir d'un langage $L \subseteq \mathcal{A}^*$ est le langage, noté \tilde{L} , constitué des miroirs des mots de L :

$$\tilde{L} = \{\tilde{\mathbf{u}} \mid \mathbf{u} \in L\}.$$

Question 1 Montrez que si L est un code, alors il en est de même de \tilde{L} .

Un mot \mathbf{v} est un *suffixe* d'un mot \mathbf{u} , s'il existe un mot \mathbf{w} tel que

$$\mathbf{u} = \mathbf{w.v}.$$

Un langage est dit *suffixe* s'il ne contient aucun mot suffixe d'un autre mot de ce langage.

Question 2 Montrez que tout langage suffixe ne contenant pas le mot vide est un code.

Question 3 À votre avis, pour quelle raison pratique les langages suffixes ne sont pas d'un usage courant pour coder un alphabet source? (Pensez au codage binaire suffixe de la source $\mathcal{S} = \{a, b, c\}$ défini par la table

a	0
b	01
c	11

et au décodage des mots $\mathbf{u}_1 = 01^{2012}$ (0 suivi de 2012 fois la lettre 1) et $\mathbf{u}_1 = 01^{2013}$ (0 suivi de 2013 fois la lettre 1).