

# Dossier : Code CLE

Nicolas Poulain

18 mars 2012

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Code CLE et entiers naturels non nuls</b>	<b>1</b>
2.1	Représentation informatique du code CLE d'un entier naturel non nul . . . . .	1
2.2	Remarques pédagogiques . . . . .	2
2.3	Premières fonctions . . . . .	2
2.4	Conversion d'un nombre en code CLE . . . . .	3
2.5	Somme de nombres en code CLE . . . . .	3
2.6	Différence (positive) de nombres en code CLE . . . . .	3
2.7	Produit de nombres en code CLE . . . . .	4
2.8	Division euclidienne de nombres en code CLE . . . . .	4
<b>3</b>	<b>Code CLE et nombres décimaux relatifs</b>	<b>4</b>
3.1	Représentation binaire en virgule flottante . . . . .	5
3.2	Représentation informatique du code CLE d'un nombre réel . . . . .	5
3.3	Remarques pédagogiques . . . . .	6
3.4	Programmation . . . . .	6

## 1 Introduction

La gestion des nombres de très grande taille peut poser des problèmes en informatique puisque généralement le type utilisé pour stocker ces entiers est de taille finie.

Pour pallier cet inconvénient, on a cherché d'autres façons d'écrire les nombres. L'une d'entre elles s'appelle le code CLE (Code à Large Echelle), elle se base sur le repérage de la position des 1 dans l'écriture binaire des nombres.

## 2 Code CLE et entiers naturels non nuls

### 2.1 Représentation informatique du code CLE d'un entier naturel non nul

Les écritures en base deux ne comportent que des 1 et des 0. Par exemple, avec un octet (8 bits) on ne peut représenter les entiers naturels que de 0 (00000000 en binaire) à 255 (11111111 en binaire). Il suffit donc de connaître la liste des positions de chiffres 1 pour connaître le nombre. La liste contient alors les exposants des puissances de deux dans l'écriture binaire, c'est-à-dire la position des bits. Ainsi, la nécessité du bit positionnel 0 disparaît. Le code CLE peut être considéré comme une formulation de l'écriture binaire. Voici deux exemples

- Le nombre 67 s’écrit 100011 en base 2. En effet on a la décomposition suivante :

$$\begin{aligned} 67 &= 1 \times 2^6 + 0 \times 2^5 + 0 \times 2^4 + 0 \times 2^3 + 0 \times 2^2 + 1 \times 2 + 1 \\ 67 &= 2^6 + 2^1 + 2^0 \end{aligned}$$

67 s’écrit alors en code CLE sous la forme  $(6; 1; 0)_{\text{CLE}}$ .

- 1548 s’écrit 11000001100 en binaire et donc  $(10; 9; 3; 2)_{\text{CLE}}$ .

L’intérêt du code CLE est évident pour les “grands nombres” dont l’écriture binaire contient peu de 1. On pense en particulier aux puissances de deux comme  $2^{64}$  dont l’écriture en code CLE est  $(64)_{\text{CLE}}$  quand son écriture décimale est 18446744073709551616.

## 2.2 Remarques pédagogiques

Dans cette première partie, on traite uniquement le cas des nombres entiers naturels non nuls afin de simplifier le problème. Par ailleurs, on se limite à la programmation de fonctions itératives afin de travailler les notions de boucles.

Ce sujet présente de nombreux avantages du point de vue mathématique et algorithmique, en effet le sujet ne nécessite que des connaissances de base d’arithmétique, cependant certains calculs ne sont pas triviaux même pour des élèves de lycée. Par ailleurs le projet est très modulaire : chaque problème se résout avec un algorithme de petite taille. De plus les difficultés sont très variées, on pourra ainsi facilement faire travailler les élèves par groupe. Les différentes méthodes et fonctions doivent cependant être cohérentes en ce qui concerne leurs entrées et sorties, ce qui oblige les groupes à bien fixer les règles du travail en commun.

Du point de vue informatique,

- Types de données. Taille des variables, de stockage et de traitement d’un nouveau type de données sont bien sûr à la base du projet. Un entier naturel en code CLE est une suite finie de nombres entiers naturels.

La première question à se poser concerne le type données le plus adapté à la représentation informatique de ces nombres en code CLE. La structure de tableau est intéressante mais dans de nombreux langages, la taille d’un tableau est fixée à la création et les opérations d’ajout ou de suppression de colonnes peuvent être difficiles.

- Choix du langage. Il s’est porté sur Python car en plus d’être libre, on trouve facilement de la documentation et que le type `list` qui est proche de la notion de code CLE est facile à manipuler puisqu’il s’agit d’une liste de longueur variable pour laquelle l’accès aux éléments peut se faire facilement par l’intermédiaire des indices.
- Programmation orientée objet. Comme le projet commence par traiter le cas des nombres entiers naturels non nuls avant de généraliser aux nombres décimaux relatifs, on fait le choix de la programmation orientée objet en commençant par définir une classe `cle` avec ses méthodes pour ensuite créer une classe dérivée `cle_etendu` qui héritera de la première et pour laquelle on pourra au besoin surcharger les méthodes.
- Commentaires et tests. Un programme Python est généralement documenté au sein même de son code source par un texte entouré de triple guillemets au début de chaque fonction afin de résumer ce qu’elle réalise et détailler les variables. Il est souvent utile de présenter des cas d’utilisation de la fonction et ce qui sera retourné, c’est pourquoi python propose un module appelé `doctest` dont le but est de permettre de réaliser des tests unitaires à l’intérieur même du commentaire. Un utilisateur pourra ainsi en regardant la documentation d’une fonction avoir directement accès à des exemples d’utilisation, exemples nécessairement à jour, puisque capables de passer les tests.

voir <http://code.google.com/p/code-cle/> pour le code à jour.

## 2.3 Premières fonctions

1. Écrire une méthode `cle_to_deci` capable de donner l’écriture décimale d’un nombre en code CLE.
2. Écrire une méthode `est_plus_grand_que` qui teste si un nombre en code CLE est plus grand qu’un autre.

## 2.4 Conversion d’un nombre en code CLE

On cherche à écrire une méthode `entier_to_cle` capable de donner l’écriture en code CLE d’un nombre entier naturel non nul. On peut procéder par divisions successives par 2. Voyons sur un exemple où on cherche l’écriture

en code CLE du nombre 13

$$\begin{array}{ll}
13 = 2 \times \underline{6} + 1 & \ell_0 \\
\underline{6} = 2 \times \underline{3} + 0 & \ell_1 \\
\underline{3} = 2 \times \underline{1} + 1 & \ell_2 \\
\underline{1} = 2 \times \underline{0} + 1 & \ell_3
\end{array}
\longrightarrow 13 = 2^3 + 2^2 + 2^0$$

## 2.5 Somme de nombres en code CLE

Avant de programmer, quelques questions :

1. Que vaut la somme  $(15)_{\text{CLE}} + (15)_{\text{CLE}}$  , puis pour  $n \in \mathbb{N}$ , la somme  $(n)_{\text{CLE}} + (n)_{\text{CLE}}$  ?
2. Que vaut la somme  $(11; 5; 0)_{\text{CLE}} + (34; 11; 5)_{\text{CLE}}$  ?

L'algorithme d'addition peut être construit de plusieurs façons, voici deux pistes

- **Somme avec tri préalable** : `somme_tri` — Concaténer puis ordonner les deux listes correspondant aux deux nombres à additionner. Ensuite, parcourir de droite à gauche à la recherche de couples  $\{n, n\}$  qu'on remplacera dans la liste par le singleton  $\{n + 1\}$ . On s'arrête quand, ayant parcouru la liste, on n'a trouvé aucune paire. Exemple

$$\begin{aligned}
(6; 5; 2; 0)_{\text{CLE}} + (5; 2)_{\text{CLE}} &= (6; 5; 5; \underline{2; 2}; 0)_{\text{CLE}} &= 2^6 + 2^5 + 2^5 + (2^2 + 2^2) + 2^0 \\
&= (6; \underline{5; 5}; 3; 0)_{\text{CLE}} &= 2^6 + (2^5 + 2^5) + 2^3 + 2^0 \\
&= (6; 6; 3; 0)_{\text{CLE}} &= (2^6 + 2^6) + 2^3 + 2^0 \\
&= (7; 3; 0)_{\text{CLE}} &= 2^7 + 2^3 + 2^0
\end{aligned}$$

- **Somme avec tri à bulle** : `somme_bulle` — Concaténer sans ordonner les deux listes en une liste  $s$  et parcourir de droite à gauche en comparant les termes consécutifs  $s(i - 1)$  et  $s(i)$  ; s'ils sont égaux, on remplace  $\{s(i - 1), s(i)\}$  par  $\{s(i - 1) + 1\}$  et si  $s(i - 1)$  est strictement inférieur à  $s(i)$ , on les permute. On s'arrête quand, ayant parcouru la liste, on n'a trouvé aucune paire ni éléments à intervertir. Exemple

$$\begin{aligned}
(6; 5; 2; 0)_{\text{CLE}} + (5; 2)_{\text{CLE}} &= (6; 5; 2; 0; 5; 2)_{\text{CLE}} &= 2^6 + 2^5 + 2^2 + 2^0 + 2^5 + 2^2 \\
&= (6; 5; 2; 5; \underline{0; 2})_{\text{CLE}} &= 2^6 + 2^5 + 2^2 + 2^5 + 2^0 + 2^2 \\
&= (6; 5; \underline{2; 5}; 2; 0)_{\text{CLE}} &= 2^6 + 2^5 + 2^2 + 2^5 + 2^2 + 2^0 \\
&= (6; 5; 5; \underline{2; 2}; 0)_{\text{CLE}} &= 2^6 + 2^5 + 2^5 + 2^2 + 2^2 + 2^0 \\
&= (6; 5; 5; 3; 0)_{\text{CLE}} &= 2^6 + 2^5 + 2^5 + 2^3 + 2^0 \\
&= (\underline{6; 6}; 3; 0)_{\text{CLE}} &= 2^6 + 2^6 + 2^3 + 2^0 \\
&= (7; 3; 0)_{\text{CLE}} &= 2^7 + 2^3 + 2^0
\end{aligned}$$

## 2.6 Différence (positive) de nombres en code CLE

La méthode `est_plus_grand_que` permet de vérifier que la différence donnera un résultat positif, on s'attache donc à construire un algorithme de soustraction du type suivant :

$$(a_1; \dots; a_n)_{\text{CLE}} - (b_1; \dots; b_m)_{\text{CLE}} \quad \text{avec} \quad \begin{cases} n \in \mathbb{N}, a_1, \dots, a_n \in \mathbb{N} \\ m \in \mathbb{N}, b_1, \dots, b_m \in \mathbb{N} \\ (a_1; \dots; a_n)_{\text{CLE}} \geq (b_1; \dots; b_m)_{\text{CLE}} \end{cases}$$

L'algorithme de différence peut être construit de plusieurs façons, voici une piste.

Remarque préliminaire : pour  $n, m \in \mathbb{N}$ ,  $n > m$ , on a  $2^n - 2^m = 2^{n-1} + \dots + 2^m$ . Aussi

$$\text{pour } a_k \geq b_1 \text{ on a } (a_k) - (b_1) = (a_k - 1, a_k - 2, \dots, b_1)$$

On souhaite calculer  $D = (a_1, \dots, a_i) - (b_1, \dots, b_j)$ . Notons  $a_k$  le plus petit élément de  $\{a_1, \dots, a_i\}$  tel que  $a_k \geq b_1$ . Alors

$$\begin{aligned} D &= (a_1, a_2, \dots, a_{k-1}, \underline{a_k}, a_{k+1}, a_{k+2}, \dots, a_i) - (\underline{b_1}, b_2, \dots, b_j) \\ &= (a_1, a_2, \dots, a_{k-1}, \underline{a_k - 1}, \underline{a_k - 2}, \dots, \underline{b_1}, a_{k+1}, a_{k+2}, \dots, a_i) - (b_2, \dots, b_j) \end{aligned}$$

Sur un exemple

$$\begin{aligned} D &= (\underline{7}, 3, 0) - (\underline{4}, 3, 1) \\ &= (\underline{6}, \underline{5}, \underline{4}, 3, 0) - (3, 1) \\ &= (6, 5, 4, \underline{3}, 0) - (\underline{3}, 1) \\ &= (6, 5, 4, 0) - (1) \\ &= (6, 5, \underline{4}, 0) - (\underline{1}) \\ &= (6, 5, \underline{3}, \underline{2}, \underline{1}, 0) \end{aligned}$$

## 2.7 Produit de nombres en code CLE

Avant de programmer, quelques questions :

1. Ecrire en code CLE le produit  $(n)_{\text{CLE}} \times (m)_{\text{CLE}}$  où  $n$  et  $m$  sont deux entiers naturels.
2. Ecrire en code CLE :
3. le produit  $(5; 2; 0)_{\text{CLE}} \times (4)_{\text{CLE}}$
4. le carré :  $(12; 4)_{\text{CLE}}^2$
5. le produit  $(5; 3)_{\text{CLE}} \times (7; 2; 1)_{\text{CLE}}$ .
6. Énoncer une règle générale.

Construire un algorithme de multiplication et le programmer.

## 2.8 Division euclidienne de nombres en code CLE

Construire un algorithme de division euclidienne pour des nombres en code CLE. Écrire les fonctions quotient et reste capables de donner le quotient et le reste dans la division de deux nombres en code CLE.

# 3 Code CLE et nombres décimaux relatifs

Afin d'aller plus loin, et de pouvoir traiter les nombres décimaux relatifs, il faut redéfinir le cadre, c'est à dire repenser la représentation du nombre. Avant cela, voyons comment est généralement défini un nombre réel dans une machine.

## 3.1 Représentation binaire en virgule flottante

La représentation en virgule flottante d'un nombre binaire sur 32 bits ou 64 bits (double précision) est définie par trois composantes :

- le signe  $S$  est représenté par un seul bit : le bit de poids fort (le plus à gauche)
- l'exposant  $E$  est codé sur les 8 bits consécutifs au signe en excès de 127 (simple précision) ou 1023 (double précision)
- la mantisse  $M$  (bits situés après la virgule) est représentée sur les 23 (simple précision) ou 55 (double précision) bits restants.

Ainsi le codage sur un mot de 32 bits se fait sous la forme suivante :

$$s \underbrace{e \dots e}_E \underbrace{m \dots m}_M$$

Dans cette représentation la valeur d'un nombre  $N$  sur 32 bit est donné par l'expression :

$$N = (-1)^s \times (1 + M \times 2^{-23}) \times 2^{(E-127)}$$

Exemple : Soit à coder la valeur 525,5.

525,5 s'écrit en base 2 de la façon suivante : 1000001101,1 on veut l'écrire sous la forme  $1,0000011011 \times 2^9$ . Par conséquent :

- le bit s vaut 1
- l'exposant  $E = 9 + 127 = 136$ , soit 10010000
- la mantisse est  $M = 0000011011$

La représentation du nombre 525,5 en binaire est :

1 10010000 000001101100000000000000

Remarque : Le zéro est le seul nombre qui ne peut pas se représenter dans le système à virgule flottante à cause de la contrainte de non nullité du premier digit ( $1, M \times 2^E$ ). Le zéro est donc représenté par une convention de représentation. La norme IEEE754 précise cette convention de représentation ( $E = M = 0$ ), c'est à dire 1 0 0.

### 3.2 Représentation informatique du code CLE d'un nombre réel

La méthode décrite ci-dessus ne permet pas d'écrire des nombres supérieurs à  $1,1111111111111111 \times 2^{127}$  de plus la précision des nombres réels non entiers est limitée. Nous allons donc étendre le système de codage CLE

1. en autorisant les éléments de la liste à prendre des valeurs négatives
2. en plaçant au premier rang un coefficient  $s \in \{-1; 0; 1\}$  qui permettra de représenter les nombres négatifs ainsi que le nombre zéro.

Ainsi, le code CLE  $(-1; 9; 3; 2; 0; -2)_{\text{CLE}}$  représente le nombre

$$-1 \times (2^9 + 2^3 + 2^2 + 2^0 + 2^{-2}) = -525,25$$

### 3.3 Remarques pédagogiques

Dans cette seconde partie, on construit la classe `cle_etendu` dérivée de la classe `cle` et on s'appuie sur le polymorphisme afin de rendre transparente la réutilisation des méthodes de la classe parent.

Dans cette partie, le typage fort du langage python ne pardonne pas les approximations. Les élèves devront veiller à toujours avoir en tête le type ou la classe des variables objets utilisés.

Des algorithmes récursifs seront mis en œuvre chaque fois que la situation s'y prêtera.

### 3.4 Programmation

Adapter (si l'héritage ne suffit pas) toutes les méthodes de la classe `cle` à la classe `cle_etendu`.