

Représentation de chaînes de caractères sous forme de nombres

Christophe Dronne

25 Octobre, 2023

Préliminaires

On définit une base quelconque, notée b , utilisant comme symboles les lettres de notre alphabet en plus d'autres symboles utilisées comme caractères typographiques. Cette base servira à représenter un nombre sous forme de texte, de la même manière que le nombre $(13303790)_{10}$ donne le nombre $(caffee)_{16}$ qui pourrait être interprété comme le mot "café".

Une implémentation possible et simple de cette base pourrait utiliser les lettres "a", "b", "c" ... "z" pour représenter les nombres 0, 1, 2, ... 25, avec $b = 26$. Cette base va être utilisée pour illustrer des exemples, mais de manière générale tout formule et propriété va être applicable à une toute base supérieure ou égale à 2.

Le but de cet exercice est de trouver une implémentation optimale des fonctions communes utilisées pour les chaînes de caractères en informatique avec ces nombres en base b .

Longueur d'un nombre

Les chaînes de caractères ont toujours une longueur associée. Pour donner un exemple, la chaîne "bonjour" a une longueur de 7, car elle est constituée de 7 caractères. Cependant, le nombre $(bonjour)_b$ n'a pas de 'longueur', donc il va falloir trouver une définition mathématique de la longueur d'un nombre.

On pose:

$\forall b \in \mathbb{N} - \{0; 1\}$, une base quelconque

$\forall n \in \mathbb{N}$, la quantité de chiffres dans k

$\forall k \in \mathbb{R}$, tel que $1 \leq k < b$

Tout entier naturel l (exclu de 0) de base b peut s'écrire tel:

$$l = k \times b^{n-1}$$

Exemple:

$534 = 5.34 \times 10^{3-1}$, avec $b = 10$, $n = 3$, $l = 534$, $k = 5.34$, suivant les conditions posées.

On pose la formule de la 'longueur' (c-à-d quantité de chiffres) d'un nombre:

$$n = E(\log_b(l)) + 1 \quad (1)$$

avec $E(x)$ la fonction de la partie entière.

Par convenance, on va noter la fonction qui attribue à un entier naturel non nul sa quantité de chiffres $len(l)$, tel que:

$$len(l) = E(\log_b(l)) + 1 \quad (2)$$

Démonstration:

On pose:

$\forall b \in \mathbb{N} - \{0; 1\}$, une base quelconque

$\forall n \in \mathbb{N}$, la quantité de chiffres dans k

$\forall k \in \mathbb{R}$, tel que $1 \leq k < b$

On a:

$$1 \leq k < b$$

k et b sont positifs, donc:

$$\Rightarrow \log_b(1) \leq \log_b(k) < \log_b(b)$$

$$\Rightarrow 0 \leq \log_b(k) < 1$$

$$\Rightarrow E(\log_b(k)) = 0$$

On reprend la formule de l :

$$l = k \times b^{n-1}$$

$$\Rightarrow \log_b(l) = \log_b(k \times b^{n-1})$$

$$\Rightarrow E(\log_b(l)) = E(\log_b(k \times b^{n-1}))$$

$$\Rightarrow E(\log_b(l)) = E(\log_b(k) + \log_b(b^{n-1}))$$

$$\Rightarrow E(\log_b(l)) = E(\log_b(k) + n - 1)$$

$(n - 1) \in \mathbb{N}$, et un entier naturel + un autre entier naturel $\in \mathbb{N}$,

donc $\forall x \in \mathbb{R}$, $E(x + n) = E(x) + n$, donc:

$$\Rightarrow E(\log_b(l)) = E(\log_b(k)) + n - 1$$

$$\Rightarrow E(\log_b(l)) + 1 = E(\log_b(k)) + n$$

or, on sait que $E(\log_b(k)) = 0$, donc:

$$\Rightarrow E(\log_b(l)) + 1 = 0 + n$$

On retrouve bien:

$$n = E(\log_b(l)) + 1$$

Indice d'un nombre