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 $(13303790)_{10}$ peut être représenté sous le nombre hexadécimal $(\text{caffee})_{16}$ qui pourrait être interprété comme le mot "caffée".

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.

Lors de cet exercice, on va souvent faire utilisation de la fonction de la partie entière, notée $E(x)$, et de la fonction de la partie fractionnaire, notée $frac(x)$, tel:

$\forall x \in \mathbb{R}$, on a:

$$E\left(x\right) \in \mathbb{N} \tag{1}$$

$$frac\left(x\right) = |x| - E\left(|x|\right) \tag{2}$$

Or, on prendra seulement en compte des valeurs positives, donc:

$$frac\left(x\right) = x - E\left(x\right) \tag{3}$$

On a les propriétés suivantes:

$0 \leqslant frac\,(x) < 1$
$\Rightarrow E\,(frac\,(x)) = 0$

$0 \leqslant x < 1$
$\Rightarrow frac\,(x) = x$
$\Rightarrow E\,(x) = E\,(frac\,(x)) = 0$

$frac\ (frac\,(x)) = frac\,(x) - E\,(frac\,(x))$
$\Rightarrow frac\,(frac\,(x)) = frac\,(x) - 0$
$\Rightarrow frac\,(frac\,(x)) = frac\,(x)$

$\forall\,(x_1; x_2) \in \mathbb{R}^2$, on a:
$E\ (x_1 + x_2) = E\,(x_1\,){+}E\,(x_2)$
si, et seulement si:
$0 \leqslant frac\,(x_1) + frac\,(x_2) < 1$
sinon:
$E\ (x_1 + x_2) = E\,(x_1\,){+}E\,(x_2) + 1$

Donc, $\forall x \in \mathbb{R}, \forall n \in \mathbb{N}$, on a:
$E\ (x + n) = E\,(x) + n$
car:
$frac\ (x) + frac\,(n) = frac\,(x) + 0 = frac\,(x)$
et
$0 \leqslant frac\,(x) < 1$
Donc $E\ (x + n) = E\,(x) + E\,(n)$
De plus, on a: E $(n) = n$, donc
$E\ (x + n)$ est bien égal à $E\,(x) + n$

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 $(\text{bonjour})_b$ 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 $1 \leqslant k < b$

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:

$$n = E\left(log_b\left(l\right)\right) + 1 \tag{4}$$

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

$$len(l) = E\left(log_b\left(l\right)\right) + 1 \tag{5}$$

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 $1 \leqslant k < b$

On a:
$1 \leqslant k < b$
k et b sont positifs, donc:
$\Rightarrow log_b\left(1\right) \leqslant log_b\left(k\right) < log_b\left(b\right)$
$\Rightarrow 0 \leqslant log_b\left(k\right) < 1$
$\Rightarrow E\left(\log_b\left(k\right)\right) = 0$

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 $\forall (n_1; n_2) \in \mathbb{N}^2$, $(n_1 + n_2) \in \mathbb{N}$,
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:

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

On a donc vérifié l'équation et justifié la validité de la fonction $len(l)$.

Exemple:
On reprend la base $b = 26$ définie dans les préliminaires:

$l = (\text{bonjour})_{26} = (481363913)_{10}$

$len\ (l) = E\,(log_{26}\,(l)) + 1$

$log_{26}\,(l) \approx 6.1361$

E $(log_{26}\,(l)) = 6$

$len\ (l) = E\,(log_{26}\,(l)) + 1 = 6 + 1$

$len\ (l) = 7$

En comptant la quantité de chiffres dans $(\text{bonjour})_{26}$, on retrouve bien 7.

Indice d'un nombre

On s'intéresse à trouver le chiffre d'indice $n \in \mathbb{N}$ d'un entier naturel quelconque l représenté en base quelconque b. On note la fonction qui permet de trouver ce chiffre $ind(l;n)$.

On pose:

$$ind(l;n) = E\left(\frac{l}{b^n}\right) - E\left(\frac{l}{b^{n+1}}\right) \times b \tag{6}$$

On note que il s'agit de l'indice **partant de la droite** à $n = 0$. Pour avoir le chiffre d'indice n **partant de la gauche**, il suffit de prendre $len(l) - n$ comme indice, soit $ind\left(l; len(l) - n\right)$, ce qui permet de suivre les conventions informatiques.

Démonstration:
On pose:
$\forall x \in \mathbb{N}$
$\forall n \in \mathbb{N}$, l'indice du nombre recherché, tel que $0 \leqslant n < len(x)$

On peut écrire x tel:

$$x = \varepsilon_1 \times b^{n+1} + c \times b^n + \varepsilon_2 \tag{7}$$

avec:
$0 \leqslant c < b$; $c \in \mathbb{N}$
$0 \leqslant \varepsilon_1$; $\varepsilon_1 \in \mathbb{N}$
$0 \leqslant \varepsilon_2 < b^n$; $\varepsilon_2 \in \mathbb{N}$

On cherche à trouver une formule de c en fonction de x, n, et b.

$$c = \frac{x - \varepsilon_1 \times b^{n+1} - \varepsilon_2}{b^n}$$

$\Rightarrow c = \frac{x-\varepsilon_2}{b^n} - \varepsilon_1 \times b$
$\Rightarrow c + \varepsilon_1 \times b = \frac{x-\varepsilon_2}{b^n}$
$\Rightarrow c + \varepsilon_1 \times b = \frac{x}{b^n} - \frac{\varepsilon_2}{b^n}$

$(c + \varepsilon_1 \times b^{n+1}) \in \mathbb{N}$, donc $c + \varepsilon_1 \times b^{n+1} = E\left(c + \varepsilon_1 \times b^{n+1}\right)$
$\Rightarrow c + \varepsilon_1 \times b^{n+1} = E\left(\frac{x}{b^n} - \frac{\varepsilon_2}{b^n}\right)$
$\Rightarrow frac\left(\frac{x}{b^n}\right) = frac\left(\frac{\varepsilon_1 \times b^{n+1} + c \times b^n + \varepsilon_2}{b^n}\right)$
$\Rightarrow frac\left(\frac{x}{b^n}\right) = frac\left(\varepsilon_1 \times b + c + \frac{\varepsilon_2}{b^n}\right)$

$(\varepsilon_1 \times b + c) \in \mathbb{N}$, donc:
$frac\ (\varepsilon_1 \times b + c) = 0$
$\Rightarrow frac\left(\varepsilon_1 \times b + c + \frac{\varepsilon_2}{b^n}\right) = frac\left(\frac{\varepsilon_2}{b^n}\right)$
$\Rightarrow frac\left(\frac{x}{b^n}\right) = frac\left(\frac{\varepsilon_2}{b^n}\right)$
$\Rightarrow frac\left(\frac{x}{b^n}\right) - frac\left(\frac{\varepsilon_2}{b^n}\right) < 1$

On reprend la formule de $(c + \varepsilon_1 \times b^{n+1})$:
c + $\varepsilon_1 \times b^{n+1} = E\left(\frac{x}{b^n} - \frac{\varepsilon_2}{b^n}\right)$
$\Rightarrow c + \varepsilon_1 \times b^{n+1} = E\left(\frac{x}{b^n}\right) - E\left(\frac{\varepsilon_2}{b^n}\right)$

Or, on a:
$0 \leqslant \varepsilon_2 < b^n$
$\Rightarrow 0 \leqslant \frac{\varepsilon_2}{b^n} < \frac{b^n}{b^n}$
$\Rightarrow 0 \leqslant \frac{\varepsilon_2}{b^n} < 1$
$\Rightarrow E\left(\frac{\varepsilon_2}{b^n}\right) = 0$

c + $\varepsilon_1 \times b^{n+1} = E\left(\frac{x}{b^n}\right) - E\left(\frac{\varepsilon_2}{b^n}\right)$
$\Rightarrow c + \varepsilon_1 \times b^{n+1} = E\left(\frac{x}{b^n}\right) - 0$
$\Rightarrow c + \varepsilon_1 \times b^{n+1} = E\left(\frac{x}{b^n}\right)$

$$c + \varepsilon_1 \times b^{n+1} = E\left(\frac{x}{b^n}\right)$$

Il suffit plus qu'à exprimer $\varepsilon_1 \times b^{n+1}$ en fonction de x, n et b.

$$\varepsilon_1 = \frac{x - c \times b^n - \varepsilon_2}{b^{n+1}}$$

$\Rightarrow \varepsilon_1 = \frac{x}{b^{n+1}} - \frac{c}{b} - \frac{\varepsilon_2}{b^{n+1}}$

$\varepsilon_1 \in \mathbb{N}$, donc $\varepsilon_1 = E\left(\varepsilon_1\right)$
$\Rightarrow \varepsilon_1 = E\left(\frac{x}{b^{n+1}} - \frac{c}{b} - \frac{\varepsilon_2}{b^{n+1}}\right)$

Or, on a:
$\frac{x}{b^{n+1}} = \frac{\varepsilon_1 \times b^{n+1} + c \times b^n + \varepsilon_2}{b^{n+1}}$
$\Rightarrow \frac{x}{b^{n+1}} = \varepsilon_1 + \frac{c}{b} + \frac{\varepsilon_2}{b^{n+1}}$

On a:

$0 \leqslant c < b$
$\Rightarrow 0 \leqslant \frac{c}{b} < 1$
$\Rightarrow E\left(\frac{c}{b}\right) = 0$

$0 \leqslant \varepsilon_2 < b^n$
$\Rightarrow 0 \leqslant \frac{\varepsilon_2}{b^{n+1}} < \frac{1}{b} < 1$
$\Rightarrow E\left(\frac{\varepsilon_2}{b^{n+1}}\right) = 0$

$\Rightarrow 0 \leqslant \frac{c}{b} + \frac{\varepsilon_2}{b^{n+1}} < 1 + \frac{1}{b} < b$
$\Rightarrow E\left(\frac{c}{b} + \frac{\varepsilon_2}{b^{n+1}}\right) = E\left(\frac{c}{b}\right) + E\left(\frac{\varepsilon_2}{b^{n+1}}\right)$

Comme dit, $\varepsilon_1 \in \mathbb{N}$, donc $frac\left(\varepsilon_1\right) = 0$ et $E\left(\varepsilon_1\right) = \varepsilon_1$
$\frac{x}{b^{n+1}} = \varepsilon_1 + \frac{c}{b} + \frac{\varepsilon_2}{b^{n+1}}$
$\Rightarrow \frac{x}{b^{n+1}} = E\left(\varepsilon_1\right) + E\left(\frac{c}{b}\right) + E\left(\frac{\varepsilon_2}{b^{n+1}}\right)$

$$\frac{x}{b^{n+1}} = E\left(\varepsilon_1\right)$$