L3I COMBINATOIRE — 9 : BASES D'ENUMERATION

Christian RONSE, ICube, Département d'Informatique de l'Université de Strasbourg

Soit b un entier > 1. Tout naturel peut se représenter *en base* b par une suite de *chiffres*, qui sont des symboles codant les entiers de 0 à $b-1$. On connaît les représentations *décimale* (en base 10) et *binaire* (en base 2), où un chiffre 0 ou 1 est appelé *bit* (contraction de *binary digit*, c.-à-d. chiffre binaire en anglais). La représentation *hexadécimale (en base 16*) est utilisée entre autres dans le codage de documents hypertexte ; ici les chiffres sont notés $0, \ldots, 9, A, \ldots, F$, où les symboles $A, \ldots, F$ correspondent aux nombres $10, \ldots, 15$ (en effet, en écrivant $10, \ldots, 15$, on représente des nombres à deux chiffres).

Pour un entier $k > 0$, et k nombres $a_0, \ldots, a_{k-1} \in \{0, \ldots, b-1\}$, la suite des chiffres codant les nombres $a_{k-1}, \ldots, a_0$ est la représentation en base b du nombre $a_{k-1}b^{k-1} + \cdots + a_0$, on écrit donc

$$[a_{k-1}, \ldots, a_0]_b = a_{k-1}b^{k-1} + \cdots + a_0 = \sum_{j=0}^{k-1} a_j b^j \ .$$

Notons que pour $k = 1$, $[a_0]_b = a_0$.

Proposition : *Soient $b > 1$ et $k > 0$. L'application $(a_{k-1}, \ldots, a_0) \mapsto [a_{k-1}, \ldots, a_0]_b$ est une bijection $\{0, \ldots, b-1\}^k \to \{0, \ldots, b^k - 1\}$. L'ordre numérique sur les $[a_{k-1}, \ldots, a_0]_b$ correspond à l'ordre lexicographique sur les $(a_{k-1}, \ldots, a_0)$, c.à-d. pour $(a_{k-1}, \ldots, a_0), (a'_{k-1}, \ldots, a'_0) \in \{0, \ldots, b-1\}^k$, on a*

$$[a_{k-1}, \ldots, a_0]_b < [a'_{k-1}, \ldots, a'_0]_b \quad \Longleftrightarrow \quad \Big(\exists i \in \{0, \ldots, k-1\},\ a_i < a'_i \text{ et } \big[\forall j > i,\ a_j = a'_j\big]\Big) \ .$$

Preuve : Pour $(a_{k-1}, \ldots, a_0) \in \{0, \ldots, b-1\}^k$, comme chaque $a_i \geq 0$, on a

$$[a_{k-1}, \ldots, a_0]_b = \sum_{j=0}^{k-1} a_j b^j \geq \sum_{j=0}^{k-1} 0 b^j = 0 \ ;$$

comme chaque $a_i \leq b-1$, on a

$$[a_{k-1}, \ldots, a_0]_b = \sum_{j=0}^{k-1} a_j b^j \leq \sum_{j=0}^{k-1} (b-1) b^j = (b-1)\Big(\sum_{j=0}^{k-1} b^j\Big) = (b-1)\Big(\frac{b^k - 1}{b-1}\Big) = b^k - 1 \ .$$

Donc $[a_{k-1}, \ldots, a_0]_b \in \{0, \ldots, b^k - 1\}$. Soient $(a_{k-1}, \ldots, a_0), (a'_{k-1}, \ldots, a'_0) \in \{0, \ldots, b-1\}^k$ tels que $(a_{k-1}, \ldots, a_0) \neq (a'_{k-1}, \ldots, a'_0)$. Comme l'ordre lexicographique sur $\{0, \ldots, b-1\}^k$ est total, on a $(a_{k-1}, \ldots, a_0) < (a'_{k-1}, \ldots, a'_0)$ ou $(a_{k-1}, \ldots, a_0) > (a'_{k-1}, \ldots, a'_0)$. Si $(a_{k-1}, \ldots, a_0) < (a'_{k-1}, \ldots, a'_0)$, cela signifie que pour un $i \in \{0, \ldots, k-1\}$ on a $a_i < a'_i$, tandis que $a_j = a'_j$ pour tout $j > i$. Donc $a'_i - a_i \geq 1$, tandis que $a'_j - a_j \geq -(b-1)$ pour $i < j$, et ainsi

$$[a'_{k-1}, \ldots, a'_0]_b - [a_{k-1}, \ldots, a_0]_b = \sum_{j=0}^{k-1} a'_j b^j - \sum_{j=0}^{k-1} a_j b^j$$

$$= \sum_{j>i} (a'_j - a_j) b^j + (a'_i - a_i) b^i + \sum_{j<i} (a'_j - a_j) b^j = (a'_i - a_i) b^i + \sum_{j<i} (a'_j - a_j) b^j$$

$$\geq 1 \cdot b^i - \sum_{j<i} (b-1) b^j = b^i - (b-1)\Big(\sum_{j<i} b^j\Big) = b^i - (b-1)\Big(\frac{b^i - 1}{b-1}\Big) = b^i - (b^i - 1) = 1 \ ,$$

par conséquent $[a_{k-1},\dots,a_0]_b < [a'_{k-1},\dots,a'_0]_b$. De même, si $(a_{k-1},\dots,a_0) > (a'_{k-1},\dots,a'_0)$, on obtient $[a_{k-1},\dots,a_0]_b > [a'_{k-1},\dots,a'_0]_b$. On en déduit que

$$[a_{k-1},\dots,a_0]_b < [a'_{k-1},\dots,a'_0]_b \quad\Longleftrightarrow\quad (a_{k-1},\dots,a_0) < (a'_{k-1},\dots,a'_0) \ ,$$

et que l'application $(a_{k-1},\dots,a_0) \mapsto [a_{k-1},\dots,a_0]_b$ est injective. Comme les deux ensembles $\{0,\dots,b-1\}^k$ et $\{0,\dots,b^k-1\}$ sont tous deux de cardinal b^k, c'est une bijection.

Notons que pour $k > 1$, $[0,a_{k-2},\dots,a_0]_b = [a_{k-2},\dots,a_0]_b$; par exemple $021 = 21$ en base 10. Donc on représentera un entier naturel avec le plus petit nombre possible de chiffres : $0 = [0]_b$ (1 chiffre), tandis que pour $n > 0$, il existe un unique $k \in \mathbb{N}$ tel que $b^{k-1} \leq n < b^k$, et on a $n = [a_{k-1},\dots,a_0]_b$ avec $a_{k-1} > 0$, en d'autres termes on représente n avec k chiffres en base b, dont celui le plus à gauche est > 0.

Pour $k > 1$ on a

$$[a_{k-1},\dots,a_0]_b = [a_{k-1},\dots,a_1]_b \times b + a_0 \ .$$

Cette identité est à la base des algorithmes d'une part pour le calcul de la représentation d'un nombre en base b (à partir de $n \in \mathbb{N}$, trouver le plus petit $k \in \mathbb{N}^*$ et $(a_{k-1},\dots,a_0) \in \{0,\dots,b-1\}^k$ tels que $[a_{k-1},\dots,a_0]_b = n$), d'autre part le calcul d'un nombre à partir de sa représentation en base b (étant donnés $k \in \mathbb{N}^*$ et $(a_{k-1},\dots,a_0) \in \{0,\dots,b-1\}^k$, donner la valeur de $[a_{k-1},\dots,a_0]_b$). Remarquons que dans la langue française, les nombres sont désignés par leur représentation en base 10, donc en pratique le problème ne se pose que pour $b \neq 10$.

Algorithme : *Représentation d'un nombre en base b.*
En entrée : n.
En sortie: k, $(a_{k-1},\dots,a_0)$.
$k := 1$;
Si $n == 0$ alors $a_0 := 0$ sinon
tant que $n \geq b$ faire
  $q := \lfloor n/b \rfloor$ et $r := n - b \cdot q$;   /quotient et reste de la division entière de n par b/
  $a_{k-1} := r$;
  k++;   /incrémenter k de 1/
  $n := q$;   [fin tant que]
$a_{k-1} := n$ [fin sinon].
Illustrons cela pour $n = 49$ et $b = 3$:
    $k = 1$, $n = 49 > 3$, $q = 16$, $r = 1$, $a_0 = 1$;
    $k = 2$, $n = 16 > 3$, $q = 5$, $r = 1$, $a_1 = 1$;
    $k = 3$, $n = 5 > 3$, $q = 1$, $r = 2$, $a_2 = 2$;
    $k = 4$, $n = 1 < 3$, $a_3 = 1$.
On a bien $[1,2,1,1]_3 = 1 \times 3^3 + 2 \times 3^2 + 1 \times 3 + 1 = 27 + 18 + 3 + 1 = 49$.

Algorithme : *Calcul d'un nombre à partir de sa représentation en base b*
En entrée : k, $(a_{k-1},\dots,a_0)$.
En sortie: n.
$n := a_{k-1}$;
Si $k > 1$ alors pour j allant de $k-2$ à 0 faire
  $n := b \cdot n + a_j$.   [fin pour]
Illustrons cela pour $[1,2,1,1]_3$:

$$
\begin{aligned}
& n = 1\,; \\
j = 2, \quad & n = 3 \times 1 + 2 = 5\,; \\
j = 1, \quad & n = 3 \times 5 + 1 = 16\,; \\
j = 0, \quad & n = 3 \times 16 + 1 = 49.
\end{aligned}
$$

On retrouve $[1,2,1,1]_3 = 49$, comme vu plus haut.

Supposons que $b = c^m$. On a alors un lien entre les représentations d'un naturel en base b et en base c. Supposons $d_{j,i} \in \{0, \dots, c-1\}$ ($j = 0, \dots, k-1$, $i = 0, \dots, m-1$), et soit

$$
a_j = [d_{j,m-1}, \dots, d_{j,0}]_c \qquad (j = 0, \dots, k-1)\ .
$$

Alors

$$
[a_{k-1}, \dots, a_0]_b = [d_{k-1,m-1}, \dots, d_{k-1,0}, \dots, d_{0,m-1}, \dots, d_{0,0}]_c\ .
$$

En effet

$$
\begin{aligned}
[a_{k-1}, \dots, a_0]_b &= \sum_{j=0}^{k-1} a_j b^j = \sum_{j=0}^{k-1} [d_{j,m-1}, \dots, d_{j,0}]_c\, b^j = \sum_{j=0}^{k-1} \Big(\sum_{i=0}^{m-1} d_{j,i} c^i \Big) (c^m)^j \\
&= \sum_{j=0}^{k-1} \sum_{i=0}^{m-1} d_{j,i} c^{mj+i} = [d_{k-1,m-1}, \dots, d_{k-1,0}, \dots, d_{0,m-1}, \dots, d_{0,0}]_c
\end{aligned}
$$

Par exemple pour $c = 2$ et $b = 16 = c^4$, $175 = [10,15]_{16}$ (175 s'écrit AF en hexadécimal), $10 = [1,0,1,0]_2$ et $15 = [1,1,1,1]_2$ (10 et 15 s'écrivent 1010 et 1111 en binaire), donc $175 = [1,0,1,0,1,1,1,1]_2$ (175 s'écrit 10101111 en binaire). Pour $b = 8 = 2^3$, on obtient alors $[1,0]_2 = 2$, $[1,0,1]_2 = 5$ et $[1,1,1]_2 = 7$, donc $[1,0,1,0,1,1,1,1]_2 = [2,5,7]_8$ (175 s'écrit 257 en base 8).

Considérons un naturel $p > 1$ et supposons que $b \equiv c$ $(mod.\ p)$. Alors pour tout $j \in \mathbb{N}$ on a $b^j \equiv c^j$ $(mod.\ p)$, donc

$$
[a_{k-1}, \dots, a_0]_b = \sum_{j=0}^{k-1} a_j b^j \equiv \sum_{j=0}^{k-1} a_j c^j \quad (mod.\ p)\ .
$$

Par exemple $b \equiv 1$ $(mod.\ b-1)$, ce qui donne

$$
[a_{k-1}, \dots, a_0]_b \equiv \sum_{j=0}^{k-1} a_j \quad (mod.\ b-1)\ .
$$

Par contre $b \equiv -1$ $(mod.\ b+1)$, donc

$$
[a_{k-1}, \dots, a_0]_b \equiv \sum_{j=0}^{k-1} (-1)^j a_j \quad (mod.\ b+1)\ .
$$

Un exemple connu est la congruence modulo 9 ou 11 d'un naturel dans sa représentation décimale. Ainsi il sera congru modulo 9 à la somme de ses chiffres, et modulo 11 à la somme de ses chiffres avec alternance de signe + et −, par exemple :

$$
4278 \equiv 4+2+7+8 = 21 \equiv 3 \quad (mod.\ 9) \qquad \text{et} \qquad 4278 \equiv -4+2-7+8 = -1 \equiv 10 \quad (mod.\ 11)\ .
$$