Représentation binaire de nombres entiers et opérations arithmétiques de base

Pierre Langlois

Représentation binaire de nombres entiers et opérations arithmétiques de base
Sujets de ce thème

- Nombres binaires non signés et signés en complément à deux.
- Addition et soustraction.
- Multiplication, division et modulo par une puissance de deux.
- Multiplication générale de nombres binaires.

Représentation binaire non signée et signée en complément à deux

- Les nombres entiers peuvent être représentés en base deux par un vecteur de bits.
- Le poids du bit d'indice *i* est 2^i.
- Dans la représentation signée en complément à deux:
  - le bit le plus significatif a un poids négatif
  - le bit le plus significatif donne le signe du nombre
- Les représentations *signe et grandeur* et *complément à un* sont moins usitées.
- En représentation non signée, la gamme de valeurs avec N bits est $0 \leq n \leq 2^N - 1$.
- En représentation signée, la gamme de valeurs avec N bits est $-2^{N-1} \leq n \leq 2^{N-1} - 1$.

b7	b6	b5	b4	b3	b2	b1	b0	valeur signée	valeur non signée
0	1	1	1	1	1	1	1	127	127
0	1	1	1	1	1	1	0	126	126
0	0	0	0	0	0	1	0	2	2
0	0	0	0	0	0	0	1	1	1
0	0	0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1	-1	255
1	1	1	1	1	1	1	0	-2	254
1	0	0	0	0	0	0	1	-127	129
1	0	0	0	0	0	0	0	-128	128

$$n_{nonsignée} = \sum_{i=0}^{N-1} b_i 2^i \qquad n_{signée} = -b_{N-1} 2^{N-1} + \sum_{i=0}^{N-2} b_i 2^i$$

Addition de nombres binaires non signés

- Il faut d'abord choisir le nombre de bits à utiliser pour représenter les deux nombres.
- L'addition se fait bit par bit.
- La somme peut nécessiter un bit de plus que celui des opérandes.

9	1001
+ 2	0010
11	1011

9	00001001
+ 182	10110110
191	10111111

9	1001
+ 7	0111
16	10000

170	10101010
+ 182	10110110
352	101100000

Changer le signe d'un nombre en complément à deux

- Pour changer le signe d'un nombre, on inverse chaque bit et on ajoute 1.
- Ceci correspond à calculer $2^N - n$, où n est le nombre dont on veux changer le signe et N est le nombre de bits pour le représenter.

3	0011
inv(3)	1100
+ 1	0001
-3	1101

-67	10111101
inv(-67)	01000010
+ 1	00000001
67	01000011

Erreur: +9 ne peut pas être représenté sur 4 bits en notation signée

9	1001
inv(9)?	0110
+ 1	0001
7	0111

9	01001
inv(9)	10110
+ 1	00001
-9	10111

Extension du signe

- Pour augmenter le nombre de bits utilisés pour représenter un nombre binaire signé, il faut faire l'extension du signe correctement.
- L'extension du signe consiste à répéter le bit de signe du nombre autant de fois que nécessaire pour obtenir le nombre de bits désirés.

+3 (4 bits)	0011
+3 (5 bits)	0 0011
+3 (6 bits)	00 0011
+3 (16 bits)	0000 0000 0000 0011

-67 (8 bits)	1011 1101
-67 (16 bits)	1111 1111 1011 1101

Addition de nombres binaires signés

- Il faut d'abord choisir le nombre de bits à utiliser pour représenter les deux nombres, en faisant l'extension du signe correctement.
- L'addition se fait bit par bit. On laisse tomber toute retenue finale.
- Il ne peut pas y avoir de débordement si les deux nombres ont un signe différent.
- Un débordement peut se produire si les deux nombres ont le même signe. On le détecte en observant que la somme a un signe différent.
- *On garantit qu'un débordement n'aura pas lieu en ajoutant un bit aux opérandes avant de faire l'addition.*

7	0111
+ (-2)	1110
5	1 0101

-6	1010
+ 7	0111
1	1 0001

7	0111
+ 7	0111
14	0 1110

Erreur:
+14 ne peut pas être représenté sur 4 bits.

7	0 0111
+ 7	0 0111
14	00 1110

OK.
+14 est représenté correctement sur 5 bits.

Soustraction de nombres binaires signés

- Une approche simple consiste à procéder comme pour l'addition, mais on change le signe du nombre à soustraire en inversant ses bits et en ajoutant 1.
- Un débordement peut avoir lieu. Les règles de détection sont similaires à celle de l'addition.
- *On garantit qu'un débordement n'aura pas lieu en ajoutant un bit aux opérandes avant de faire la soustraction.*

7	0111
- 2: inv(2)	1101
	1
5	1 0101

7	0111
- (-2): inv(-2)	0001
	1
9	0 1001

Comparaison de la grandeur de nombres binaires signés

- La comparaison de grandeur se fait par une simple soustraction.
- Le signe de la différence indique lequel des opérandes est le plus grand
  - Signe positif: le premier opérande est plus grand;
  - Signe négatif: le second opérande est plus grand;
  - Résultat zéro: les opérandes sont égaux.
- On combine ensuite le signe (et la valeur zéro) de la différence avec l'opérateur de comparaison pour produire la valeur booléenne de la comparaison.

7 <? -5
7
- (-5)
12

12 > 0, donc 7 > -5.
L'opérateur est 'plus petit'.
Donc le résultat est FAUX.

7 >=? 7
7
- 7
0

7 – 7 == 0, donc 7 == 7.
L'opérateur est 'plus grand ou égal'.
Donc le résultat est VRAI.

Multiplication et division par une puissance de deux

- Dans toute base, la multiplication par une puissance de la base revient à décaler les chiffres vers la gauche.
- La division par une puissance de la base revient à décaler les chiffres vers la droite.
- En arithmétique entière, pour la division on ne garde que la partie entière du résultat.

$$37 \times 100 = 3700$$
$$5 \times 1000 = 5000$$

$$\lfloor 2340 \div 10 \rfloor = 234$$
$$\lfloor 2340 \div 1000 \rfloor = 2$$

$$5 \times 2^2 = 5 \times 4 = 20$$
$$5 \times 2^4 = 5 \times 16 = 80$$

$$0101_2 \times 2^2 = 01\ 0100_2 = 20$$
$$0101_2 \times 2^4 = 0101\ 0000_2 = 80$$

$$\lfloor 46 \div 2^1 \rfloor = \lfloor 46 \div 2 \rfloor = 23$$
$$\lfloor 46 \div 2^2 \rfloor = \lfloor 46 \div 4 \rfloor = 11$$

$$\lfloor 101110_2 \div 2^1 \rfloor = 10111_2 = 23$$
$$\lfloor 101110_2 \div 2^2 \rfloor = 1011_2 = 11$$

Modulo par une puissance de deux

- Dans toute base, le modulo par une puissance de la base revient à extraire des chiffres de poids faible.
- Le nombre de chiffres extraits correspond à la puissance de la base de l'opérande du modulo.

$$3758 \% 10 = 3758 \% 10^1 = 8$$
$$3758 \% 100 = 3758 \% 10^2 = 58$$
$$3758 \% 1000 = 3758 \% 10^3 = 758$$

$$247 \% 4 = 247 \% 2^2 = 3$$
$$247\% 16 = 247 \% 2^4 = 7$$
$$247 \% 64 = 247 \% 2^6 = 55$$

$$1111\ 0111_2 \% 2^2 = 11_2 = 3$$
$$1111\ 0111_2 \% 2^4 = 0111_2 = 7$$
$$1111\ 0111_2 \% 2^6 = 11\ 0111_2 = 55$$

Multiplication générale de nombres binaires

- La multiplication binaire se fait comme pour la base 10.
- Pour les nombres non signés, le nombre de bits requis pour représenter le produit est égal à la somme du nombre de bits des opérandes.
- Pour les nombres signés, il faut 1 bit de moins que pour le cas non signé.

7	0111
× 5	0101
	0111 00000 011100 0000000
35	0100011

Représentation	Grandeur maximale (*N* bits)	Produit maximal	Exemple avec *N* = 3 bits
Non signée	$2^N - 1$	$(2^N - 1) \times (2^N - 1) = 2^{2N} - 2 \times 2^N + 1$ Il faut $2N$ bits	Valeur max: 7 Produit max: 49 (6 bits)
Signée	-2^{N-1}	$(-2^{N-1}) \times (-2^{N-1}) = 2^{2N-2}$ Il faut $2N - 1$ bits	Valeur max: -4 Produit max: +16 (5 bits)

Vous devriez maintenant être capable de …

- Utiliser la représentation binaire de nombres entiers signés et non signés dans des problèmes numériques. (B3)
- Effectuer les opérations arithmétiques suivantes sur des entiers en représentation binaire, en tenant compte du nombre de bits des opérandes et du résultat:
  - addition et soustraction;
  - comparaison;
  - multiplication, division et modulo par une puissance de deux;
  - multiplication générale. (B3)

Code	Niveau (http://fr.wikipedia.org/wiki/Taxonomie_de_Bloom)
B1	Connaissance - mémoriser de l'information.
B2	Compréhension – interpréter l'information.
B3	Application – confronter les connaissances à des cas pratiques simples.
B4	Analyse – décomposer un problème, cas pratiques plus complexes.
B5	Synthèse – expression personnelle, cas pratiques plus complexes.