BROUILLON - CANDIDAT - QUELS DÉCALAGES ENTRE LES ENTIERS SIGNÉS ET LES NON SIGNÉS ?

CHRISTOPHE BAL

Document, avec son source LaTeX, disponible sur la page
https://github.com/bc-writing/drafts.

Mentions « légales »

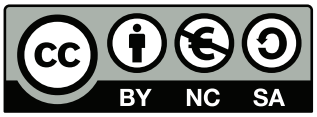

Table des matières

1. Entiers non signés

Dans ce document, nous allons travailler uniquement avec des entiers non signés, c'est à dire des naturels dont l'écriture en base 2 tient sur 8 bits même si les raisonnements tenus se généralisent à 16, 32 ou 64 bits. Ce choix tient juste à des contraintes de lisibilité du document.

Voici des exemples de représentations de type little-endian : on part à gauche de la puissance la plus haute, pour nous ce sera toujours 2^7, puis on va vers la plus basse *(c'est le même principe que l'écriture des naturels en base* 10*)*.

1	**0**	**0**	**0**	**0**	**1**	**0**	**1**

$\leftarrow 133 = 128 + 4 + 1 = 2^7 + 2^2 + 2^0$

0	**0**	**1**	**0**	**0**	**1**	**1**	**0**

$\leftarrow 38 = 32 + 4 + 2 = 2^5 + 2^2 + 2^1$

Avec 8 bits on peut représenter les naturels de 0 à $1 + 2 + 2^2 + \cdots + 2^7 = 2^8 - 1 = 255$.

Date: 17 Mars 2019 - 11 Janvier 2020.

2. Entiers signés

2.1. **Comment ça marche ?** Voici des exemples de représentations toujours sur 8 bits et de type little-endian où la case rouge indique le bit utilisé pour coder le signe.

0	**1**	**1**	**1**	**1**	**0**	**0**	**0**	$\leftarrow 120 = 64+32+16+8 = \boxed{0\times(-2^7)} + 2^6+2^5+2^4+2^3$
1	**0**	**0**	**0**	**1**	**0**	**0**	**0**	$\leftarrow -120 = -128+8 = \boxed{1\times(-2^7)} + 2^3$
1	**0**	**0**	**0**	**0**	**0**	**0**	**0**	$\leftarrow -128 = \boxed{1\times(-2^7)} + 0$

Avec 8 bits on peut représenter les naturels de $-2^7 = -128$ à $1+2+2^2+\cdots+2^6 = 2^7-1 = 127$.

Ce choix permet d'effectuer des additions relatives toujours sous forme d'addition bit à bit, et donc par conséquent de faire aussi des soustractions via des additions bit à bit. Voici quelques exemples sans dépassement de capacité.

- Addition de deux positifs.

	0	**1**	**0**	**1**	**0**	**1**	**0**	**0**	$\leftarrow 84 = 64+16+4$
$+$	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**1**	$\leftarrow 17 = 16+1$
$=$	**0**	**1**	**1**	**0**	**0**	**1**	**0**	**1**	$\leftarrow 101 = 64+32+4+1$

- Addition de deux négatifs *(la retenue finale hors capacité est ignorée)*.

	1	**0**	**1**	**0**	**1**	**1**	**0**	**0**	$\leftarrow -84 = \boxed{-128} + 44 = \boxed{-128} + 32+8+4$
$+$	**1**	**1**	**1**	**0**	**1**	**1**	**1**	**1**	$\leftarrow -17 = \boxed{-128} + 111 = \boxed{-128} + 64+32+8+4+2+1$
$=$	**1**	**0**	**0**	**1**	**1**	**0**	**1**	**1**	$\leftarrow -101 = \boxed{-128} + 27 = \boxed{-128} + 16+8+2+1$

- Addition de deux entiers de signes différents.

	0	**1**	**0**	**1**	**0**	**1**	**0**	**0**	$\leftarrow 84 = 64+16+4$
$+$	**1**	**0**	**0**	**1**	**1**	**0**	**1**	**1**	$\leftarrow -101 = \boxed{-128} + 27 = 16+8+2+1$
$=$	**1**	**1**	**1**	**0**	**1**	**1**	**1**	**1**	$\leftarrow -17 = \boxed{-128} + 111 = -128+64+32+8+4+2+1$

- Autre addition de deux entiers de signes différents *(la retenue finale hors capacité est ignorée)*.

	0	**1**	**0**	**1**	**0**	**1**	**0**	**0**	$\leftarrow 84 = 64+16+4$
$+$	**1**	**1**	**0**	**0**	**0**	**0**	**1**	**1**	$\leftarrow -61 = \boxed{-128} + 67 = 64+2+1$
$=$	**0**	**0**	**0**	**1**	**0**	**1**	**1**	**1**	$\leftarrow 23 = 16+4+2+1$

2.2. **Qu'est-ce qui motive ces choix ?** Pour comprendre comment découvrir ce type de procédé, nous allons raisonner en base 10 et imaginer que nous voulions calculer $189-32 = 157$ en utilisant uniquement des additions. Tout un chacun sait effectuer ce calcul comme suit.

	1	8	9
$-$		3	2
$=$	1	5	7

Nous décidons d'ajouter 1000 à $189-32$. Nous obtenons :

		1	8	9
$-$			3	2
$+$	1	0	0	0
$=$	1	1	5	7

Comme le résultat initial tenait sur les trois chiffres de droite, l'ajout de 1000 donne un nouveau résultat dont nous savons que les trois chiffres de droite donnent le résultat de $189 - 32$ car la soustraction $1157 - 1000$ revient à ignorer le 1 tout à droite.

D'autre part, $1000 + 189 - 32 = 189 + 1000 - 32 = 189 + 968$ est une simple addition.

Notons enfin que $1000 - 32 = 1 + 999 - 032 = 1 + 967$ où 9 , 6 et 7 sont les compléments à 9 de 0 , 3 et 2 respectivement.

En résumé, nous avons fait comme suit où • indique une ligne où a été fait le calcul d'un complément à 9 plus 1. Nous avons encerclé le signe final pour le différencier des signes d'opération :

$$\begin{array}{cccccc}
 & & 1 & 8 & 9 \\
- & & 0 & 3 & 2 \\
\hline\hline
 & & 1 & 8 & 9 \\
+ & \bullet & 9 & 6 & 8 \\
\hline
= & 1 & 1 & 5 & 7 \\
\hline\hline
 & \boxed{+} & 1 & 5 & 7
\end{array}$$

Très bien mais que se passe-t-il avec $32 - 189 = -157$ qui est négatif ? Testons pour voir.

$$\begin{array}{cccccc}
 & & & 3 & 2 \\
- & & 1 & 8 & 9 \\
\hline\hline
 & & & 3 & 2 \\
+ & \bullet & 8 & 1 & 1 \\
\hline
= & 0 & 8 & 4 & 3
\end{array}$$

Oups ! Nous ne voyons pas directement 157. Normal ! Il reste à effectuer $843 - 1000 = -157$. Dans un tel cas, on note que $843 - 1000 = -(1000 - 843)$ s'obtient via un complément à 9 plus 1 de 843 tout en ajoutant un signe moins. En résumé, nous procédons comme suit en notant qu'ici nous avons un 0 tout à gauche juste avant le résultat final contrairement au 1$^{\text{er}}$cas où il y avait un 1.

$$\begin{array}{cccccc}
 & & & 3 & 2 \\
- & & 1 & 8 & 9 \\
\hline\hline
 & & & 3 & 2 \\
+ & \bullet & 8 & 1 & 1 \\
\hline
= & 0 & 8 & 4 & 3 \\
\hline\hline
 & \boxed{-} & 1 & 5 & 7
\end{array}$$

Ne nous emballons pas trop vite car il reste un cas problématique que nous n'avons pas abordé. Examinons ce qu'il se passe avec $-32 - 189 = -221$.

$$\begin{array}{cccccc}
 & \boxed{-} & 0 & 3 & 2 \\
- & & 1 & 8 & 9 \\
\hline\hline
 & \bullet & 9 & 6 & 8 \\
+ & \bullet & 8 & 1 & 1 \\
\hline
= & 1 & 7 & 7 & 9
\end{array}$$

Nous devons affiner notre règle via la valeur du chiffre le plus à gauche car sinon ici nous aurions un résultat positif ! En fait, il est facile de voir qu'il suffit de tenir compte du nombre de compléments à 9 plus 1 initiaux effectués [1] au chiffre tout à gauche qui sera 0 ou 1. En effet, en notant n le nombre de compléments à 9 plus 1 effectués avant l'étape finale et g le chiffre à gauche de l'avant dernier résultat, nous avons les faits suivants.

1. C'est à dire le nombre de 1000 ajoutés.

(1) **Cas 1 :** pour $189 - 32$, $n = 1$, $g = 1$ et nous avions juste lu le résultat en ignorant le chiffre tout à gauche sans faire de complément à 9 plus 1 supplémentaire.

(2) **Cas 2 :** pour $32 - 189$, $n = 1$, $g = 0$ et nous avions fait un complément à 9 plus 1 supplémentaire, lu le résultat en ignorant le chiffre tout à gauche puis enfin ajouter un signe moins.

(3) **Cas 3 actuel :** pour $-32 - 189$, $n = 2$, $g = 1$ et nous devons faire comme dans le 2e cas, ceci nous donnant :

$$\begin{array}{rcccc} & \boxed{-} & 0 & 3 & 2 \\ - & & 1 & 8 & 9 \\ \hline\hline & \bullet & 9 & 6 & 8 \\ + & \bullet & 8 & 1 & 1 \\ \hline = & 1 & 7 & 7 & 9 \\ \hline\hline & \boxed{-} & 2 & 2 & 1 \end{array}$$

Nous avons donc envie de dire que $n - g \in \{0\,;1\}$ donne le nombre de compléments à un supplémentaire à faire ainsi que le nombre de signe moins à ajouter après avoir lu le résultat en ignorant le chiffre tout à gauche. Que c'est beau !

Les deux cas suivants achèvent notre exploration expérimentale en confirmant notre conjecture.

$$\begin{array}{rcccc} & & 0 & 3 & 2 \\ + & & 1 & 8 & 9 \\ \hline = & & 2 & 2 & 1 \\ \hline\hline & \boxed{+} & 2 & 2 & 1 \end{array}$$

$$\begin{array}{rcccc} & \boxed{-} & 0 & 3 & 2 \\ + & & 1 & 8 & 9 \\ \hline\hline & \bullet & 9 & 6 & 8 \\ + & & 1 & 8 & 9 \\ \hline = & 1 & 1 & 5 & 7 \\ \hline\hline & \boxed{+} & 1 & 5 & 7 \end{array}$$

Exercice 1. *Dans la méthode ci-dessous, il existe des cas problématiques. Lesquels ?*

Exercice 2. *En laissant les cas problématiques de côté, démontrer le caractère général de la méthode que nous avons juste exposée via quelques exemples.*

2.3. **Pourquoi ça marche ?** Revenons à nos entiers signés écrits sur 8 bits en reprenant le calcul de $84 - 101 = -17$ tout en nous inspirant de ce qui a été vu dans la section précédente avec la base 10. Ici 128 joue le même rôle que 1000 avant. Nous avons ajouté des cases grises pour le bit du calcul intermédiaire, et nous avons indiqué les complémentations en mettant sur fond vert les bits les plus à gauche..

$$\begin{array}{rcccccccc} & \mathbf{0} & & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ - & \mathbf{0} & & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} \\ \hline\hline & \mathbf{0} & & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ + & \mathbf{1} & & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{1} \\ \hline = & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \hline\hline - & \mathbf{0} & & \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{array}$$

La dernière complémentation avec ajout d'un signe vient du O sur fond gris et du fait q'une seule complémentation préparatoire a été faite avant. Tout s'éclaire !

Reprenons de même le cas de $84 - 61 = 23$ où l'on avait une retenue à ignorer. On fait comme précédemment mais en utilisant que les 8 bits disponibles.

	0		1	0	1	0	1	0	0
−	0		0	1	1	1	1	0	1
	0		1	0	1	0	1	0	0
+	1		1	0	0	0	0	1	1
=	1	1	0	0	1	0	1	1	1
+	0		0	0	1	0	1	1	1

Avec le nouvel éclairage, nous n'avons qu'à garder les bits sur fond blanc, le signe du résultat étant positif [2].

Dans les deux cas précédents, au lieu de considérer un bit intermédiaire, il suffit d'effectuer directement le calcul de retenue sur le bit de gauche modulo deux, ce calcul traduisant celui sur le nombre de complémentations préparatoires et du chiffre intermédiaire en gris dans nos deux exemples ci-dessus. On peut ainsi faire plus directement comme suit.

	0	1	0	1	0	1	0	0
−	0	1	1	0	0	1	0	1
	0	1	0	1	0	1	0	0
+	1	0	0	1	1	0	1	1
=	1	1	1	0	1	1	1	1
−	0	0	0	1	0	0	0	1

	0	1	0	1	0	1	0	0
−	0	0	1	1	1	1	0	1
	0	1	0	1	0	1	0	0
+	1	1	0	0	0	0	1	1
=	0	0	0	1	0	1	1	1
+	0	0	0	1	0	1	1	1

Plus généralement, avec la façon de stocker les entiers signés, nous avons alors les correspondances suivantes où les entiers additionnés a et b sont dans $[-128\,;127]$ et leur somme aussi, ce qui revient à ne considérer que les cas de non dépassement de capacité.

	0	⋯
−	0	⋯
	0	⋯
+	1	⋯
=	0	⋯

Soustraction 1 de deux naturels

	0	⋯
−	0	⋯
	0	⋯
+	1	⋯
=	1	⋯

Soustraction 2 de deux naturels

−	0	⋯
−	0	⋯
	1	⋯
+	1	⋯
=	0	⋯

Addition 1 de deux relatifs négatifs

−	0	⋯
−	0	⋯
	1	⋯
+	1	⋯
=	1	⋯

Addition 2 de deux relatifs négatifs

Les soustractions et l'addition 2 ne posent aucun souci. Par contre l'addition 1 est problématique avec son résultat positif. En fait cette addition contredit notre hypothèse de non dépassement de capacité. Nous allons voir pourquoi.

2. Il y a eu une seule complémentation préparatoire et le bit intermédiaire est 1.

Le cas de l'addition 1 correspond à $(a\,;b) \in [-128\,;-1]^2$ tel que $(128+a)+(128+b) \in [0\,;127]$ soit $0 \leq 256+a+b \leq 127$ *i.e.* $-256 \leq a+b \leq -129$ ce qui correspond à un dépassement de capacité comme annoncé.

Ceci achève de démontrer la validité des procédés d'addition et de soustraction d'entiers signés dans les cas de non dépassement de capacité. Notez que le cas évident d'une addition de deux naturels a été omis, et aussi que $-a+b = b-a = b+(-a)$ et le fait qu'un changement de signe n'est autre qu'un complément à 1 plus 1 permettent de compléter les cas non indiqués ci-dessus.

Exercice 3. *Étudiez plus généralement le cas d'une base $b \in \mathbb{N}_{\geq 3}$ quelconque.*

3. Division par 2 et décalage à droite

3.1. **Entiers non signés.** Lorsque l'on calcule $(n \div 2)$ le quotient entier par défaut d'un naturel non signé n divisé par 2, il suffit de retirer son bit de poids faible. Voici un exemple.

1	**1**	**0**	**1**	**0**	**1**	**1**	**1**	$\leftarrow n_0 = 128+64+16+4+2+1 = 215$
0	**1**	**1**	**0**	**1**	**0**	**1**	**1**	$\leftarrow n_1 = n_0 \div 2 = 107$
0	**0**	**1**	**1**	**0**	**1**	**0**	**1**	$\leftarrow n_2 = n_1 \div 2 = 53$
0	**0**	**0**	**1**	**1**	**0**	**1**	**0**	$\leftarrow n_3 = n_2 \div 2 = 26$

Chaque étape correspond à un décalage vers la droite tout en complétant par un zéro à gauche.

3.2. **Entiers signés.** Examinons ce qu'il se passe lorsque l'on calcule $(n \div 2)$ le quotient entier par défaut d'un naturel signé $n<0$ divisé par 2.

1	**1**	**0**	**1**	**0**	**1**	**1**	**1**	$\leftarrow n_0 = -128+64+16+4+2+1 = -41$
1	**1**	**1**	**0**	**1**	**0**	**1**	**1**	$\leftarrow n_1 = n_0 \div 2 = -21 = -128+64+32+8+2+1$

On en déduit que quelque soit le signe de l'entier signé, le quotient entier par défaut correspond à un décalage vers la droite tout en conservant le bit de signe tout à gauche. Notez bien que l'on décale tous les bits y compris celui du signe. Démontrons nos affirmations.

Le cas $n \in [0\,;127]$ étant évident, considérons $n \in [-128\,;-1]$. La représentation de n s'obtient alors via $n = -2^7 + \displaystyle\sum_{0\leq k\leq 6} b_k\, 2^k$ avec chaque b_k dans $\{0\,;1\}$. Nous avons alors :

$$n \div 2 = -2^6 + \sum_{1\leq k\leq 6} b_k\, 2^{k-1}$$

$$n \div 2 = -2\times 2^6 + 2^6 + \sum_{0\leq k\leq 5} b_{k-1}\, 2^k$$

$$n \div 2 = -2^7 + 2^6 + \sum_{0\leq k\leq 5} b_{k-1}\, 2^k$$

La dernière identité correspond bien au décalage avec conservation de signe qui a été annoncé plus haut.

3.3. **Synthèse.** Nous constatons que nous avons deux types de décalage à droite. Dans le cas des entiers non signés, le décalage vers la droite est noté `>>>` et il est appelé *« décalage logique à droite »*. Dans le cas des entiers signés, le décalage vers la droite est noté `>>` et il est appelé *« décalage arithmétique à droite »*.

4. Multiplication par 2 et décalage à gauche

4.1. **Entiers non signés.** Lorsque l'on calcule $2n$ le double d'un naturel non signé n, il suffit de retirer son bit de poids fort tout en ajoutant un zéro comme nouveau bit de poids faible. Voici un exemple sans dépassement de capacité.

0	**0**	**0**	**1**	**0**	**1**	**1**	**1**	$\leftarrow n_0 = 16+4+2+1 = 23$
0	**0**	**1**	**0**	**1**	**1**	**1**	**0**	$\leftarrow n_1 = 2n_0 = 46$
0	**1**	**0**	**1**	**1**	**1**	**0**	**0**	$\leftarrow n_2 = 2n_1 = 92$
1	**0**	**1**	**1**	**1**	**0**	**0**	**0**	$\leftarrow n_3 = 2n_2 = 184$

Chaque étape correspond à un décalage vers la gauche tout en complétant par un zéro à droite.

4.2. **Entiers signés.** Examinons ce qu'il se passe lorsque l'on calcule $2n$ si $n < 0$ est un naturel signé. Voici un exemple sans dépassement de capacité.

1	**1**	**0**	**1**	**0**	**1**	**1**	**1**	$\leftarrow n_0 = -128+64+16+4+2+1 = -41$
1	**0**	**1**	**0**	**1**	**1**	**1**	**0**	$\leftarrow n_1 = 2n_0 = -82 = -128+32+8+4+2$

On effectue un décalage à gauche de tous les bits, y compris celui du signe, comme avec les entiers non signés. Mais est-ce toujours vrai ? Que se passe-t-il si le 2^{e} bit à gauche est nul ? En fait, sous l'hypothèse de non dépassement de capacité, tout fonctionne sans souci comme nous allons l'expliquer proprement tout de suite.

Le cas $n \in [0\,;127]$ étant évident, considérons $n \in [-128\,;-1]$. La représentation de n s'obtient alors via $n = -2^7 + \sum_{0\leq k\leq 6} b_k\, 2^k$ avec chaque b_k dans $\{0\,;1\}$. Nous avons alors :

$$2n = -2^8 + \sum_{0\leq k\leq 6} b_k\, 2^{k+1}$$

$$2n = -2^8 + b_6\, 2^7 + \sum_{1\leq k\leq 6} b_{k-1}\, 2^k$$

$$2n = (b_6 - 2)\, 2^7 + \sum_{1\leq k\leq 6} b_{k-1}\, 2^k$$

Si $b_6 = 1$, nous avons bien le décalage annoncé. Supposons donc que $b_6 = 0$. Dans ce cas, en notant $n = -2^7 + S$, nous obtenons :

$$S = \sum_{0\leq k\leq 6} b_k\, 2^k$$

$$S = \sum_{0\leq k\leq 5} b_k\, 2^k$$

$$S \leq \sum_{0\leq k\leq 5} 2^k$$

$$S \leq 2^6 - 1$$

Nous en déduisons que $n = -2^7 + S \leq -2^7 + 2^6 - 1 = -2^6 - 1$, puis ensuite $2n \leq -2^7 - 2 < -2^7$ ce qui est bien un cas de dépassement de capacité.

4.3. **Synthèse.** Dans le cas des entiers signés ou non, nous utilisons le même décalage vers la gauche qui est noté << . Pour préciser sur quel type d'entiers on effectue un décalage, on parlera de *« décalage logique à gauche »* pour les entiers non signés, et de *« décalage arithmétique à gauche »* pour les entiers signés.