Exercice 1. Entiers signés

Question 1.a – Soit $n = 80$.

$$n \geqslant 0 \qquad \text{donc} \qquad \mathrm{RCA2}(n) = \mathrm{RB2}_{64}(n)$$

. Or :

$$n = 80 = 64 + 16 = (100\ 0000)_2 + (1\ 0000)_2 = (101\ 0000)_2$$

$$\boxed{\mathrm{RCA2}(80) = (\underbrace{0\ldots0}_{\text{57 zéros}}\ 101\ 0000).}$$

Question 1.b – Soit $n = -100$.

$$n < 0 \qquad \text{donc} \qquad \mathrm{RCA2}(n) = \mathrm{Flip}(\mathrm{RB2}_{64}(-n-1))$$

Or :

$$-n - 1 = 99 = 64 + 32 + 2 + 1 = (110\ 0011)_2 = (\underbrace{0\ldots0}_{\text{57 zéros}}\ 110\ 0011)_2$$

Donc :

$$\boxed{\mathrm{RCA2}(-100) = (\underbrace{1\ldots1}_{\text{57 uns}}\ 001\ 1100)_2.}$$

Question 2.a – Soit n l'entier tel que $\mathrm{RCA2}(n) = (\underbrace{0\ldots0}_{\text{57 zéros}}\ 100\ 1001)$. Comme le bit de poids fort de $\mathrm{RCA2}(n)$ est 0, on a :

$$n \geqslant 0 \qquad \text{et} \qquad \mathrm{RCA2}(n) = \mathrm{RB2}_{64}(n)$$

Donc :

$$n = (100\ 1001)_2 = 1 + 8 + 64 = 73$$

$$\boxed{(\underbrace{0\ldots0}_{\text{57 zéros}}\ 100\ 1001) \text{ est la représentation complément à deux de } 73.}$$

Question 2.b – Soit n l'entier tel que $\mathrm{RCA2}(n) = (\underbrace{1\ldots1}_{\text{58 uns}}\ 00\ 1001)$. Comme le bit de poids fort de $\mathrm{RCA2}(n)$ est 1, on a :

$$n < 0 \qquad \text{et} \qquad \mathrm{RCA2}(n) = \mathrm{Flip}(\mathrm{RB2}_{64}(-n-1))$$

Donc :

$$\mathrm{RB2}_{64}(-n-1) = (\underbrace{0\ldots0}_{\text{58 zéros}}\ 11\ 0110)$$

Donc :

$$-n - 1 = (11\ 0110)_2 = 2 + 4 + 16 + 32 = 54$$

$$\boxed{(\underbrace{1\ldots1}_{\text{58 uns}}\ 00\ 1001) \text{ est la représentation complément à deux de } -55.}$$

Exercice 2. Conversion de réels

Question 1.a – Soit $x = -7.6875$, alors :

$$x = -(4+2+1+0.5+0.125+0.0625) = -(111.1011)_2 = (-1)^1(1.1110\ 11)_2 \times 2^2$$

Soit $e = 2$, alors :

$$\mathrm{RBZ}(e) = \mathrm{RB2}_{11}(e + 2^{10} - 1) = \mathrm{RB2}_{11}(2^{10}+1) = (1\underbrace{0\dots0}_{9\text{ zéros}}1)$$

Finalement :

$$\boxed{\mathrm{RB64}(-7.6875) = (1 \mid 1\underbrace{0\dots0}_{9\text{ zéros}}1 \mid 1110\ 11\ \underbrace{0\dots0}_{46\text{ zéros}})}$$

Question 1.b – Soit $x = -2^{-1022}$, alors :

$$x = (-1)^0 \times (1.0)_2 \times 2^{-1022}$$

Soit $e = 2$, alors :

$$\mathrm{RBZ}(e) = \mathrm{RB2}_{11}(e + 2^{10} - 1) = \mathrm{RB2}_{11}(e + 1024 - 1) = \mathrm{RB2}_{11}(1) = (\underbrace{0\dots0}_{10\text{ zéros}}\ 1)$$

Finalement :

$$\boxed{\mathrm{RB64}(2^{-1022}) = (0 \mid \underbrace{0\dots0}_{10\text{ zéros}}\ 1 \mid \underbrace{0\dots0}_{52\text{ zéros}})}$$

Question 2.a – Soit x tel que $\mathrm{RB64}(x) = (0 \mid 1\underbrace{0\dots0}_{8\text{ zéros}}10 \mid 1101\ \underbrace{0\dots0}_{48\text{ zéros}})$ et e tel que $\mathrm{RBZ}(e) = (1\underbrace{0\dots0}_{8\text{ zéros}}10)$.

Alors :

$$\mathrm{RBZ}(e) = \mathrm{RB2}_{11}(e + 2^{10} - 1) \qquad \text{donc} \qquad e + 2^{10} - 1 = (1\underbrace{0\dots0}_{8\text{ zéros}}10)_2 = 2^{10} + 2$$

On a donc $e = 3$ et :

$$x = (-1)^0(1.1101)_2 \times 2^3 = (1110.1)_2 = 2+4+8+0.5 = 14.5$$

$$\boxed{(0 \mid 1\underbrace{0\dots0}_{8\text{ zéros}}10 \mid 1101\ \underbrace{0\dots0}_{48\text{ zéros}}) \text{ est la représentation } \texttt{binary64} \text{ de } 14.5}$$

Question 2.b – Soit x tel que $\mathrm{RB64}(x) = (1 \mid 0\underbrace{1\dots1}_{9\text{ uns}}0 \mid 01\ \underbrace{0\dots0}_{50\text{ zéros}})$ et e tel que $\mathrm{RBZ}(e) = (0\underbrace{1\dots1}_{9\text{ uns}}0)$.

Alors :

$$\mathrm{RBZ}(e) = \mathrm{RB2}_{11}(e + 2^{10} - 1) \qquad \text{donc} \qquad e + 2^{10} - 1 = (0\underbrace{1\dots1}_{9\text{ uns}}0)_2$$

Ainsi :

$$-e = (1\ \underbrace{0\dots0}_{10\text{ zéros}})_2 - (0\underbrace{1\dots1}_{9\text{ uns}}0)_2 - 1 = (10)_2 - 1 = 1.$$

On a donc $e = -1$ et :

$$x = (-1)^1(1.01)_2 \times 2^{-1} = -(0.101)_2 = -(0.5 + 0.125) = -0.625$$

$$\boxed{(1 \mid 0\underbrace{1\dots1}_{9\text{ uns}}0 \mid 01\ \underbrace{0\dots0}_{50\text{ zéros}}) \text{ est la représentation } \texttt{binary64} \text{ de } -0.625}$$

Exercice 5.

Question 1 – Le plus petit flottant strictement supérieur à 1 est le flottant représenté par les bits $(0 \mid 0\ \underbrace{1\ldots1}_{\text{10 uns}} \mid \underbrace{0\ldots0}_{\text{51 zéros}}1)$ qui vaut :

$$1 + 2^{-52} \approx 1 + 2.2 \times 10^{-16}.$$

Question 2 – Soit $x = (10.1\ 110\ 110\ 110\ \ldots)_2$, alors :

$$x = 2.5 + (0.0\ 110\ 110\ 110\ \ldots)_2$$

Posons $y = x - 2.5$, alors :

$$\begin{aligned}
&y = (0.0\ 110\ 110\ 110\ \ldots)_2 \\
&y \times 2^4 = (110.110\ 110\ \ldots)_2 \\
&y \times 2^4 - 6 = (0.110\ 110\ \ldots)_2 = 2y \\
&14y = 6
\end{aligned}$$

Donc $y = \dfrac{3}{7}$ et $\boxed{x = 2 + \frac{13}{14}}$

Question 3 – Soit $x = 9.55$. Alors :

$$\begin{aligned}
&2^0 \times x = (1001)_2 + 0.55 \\
&2^1 \times x = (1\ 0010)_2 + 1.1 = (1\ 0011)_2 + 0.1 \\
&2^2 \times x = (10\ 0110)_2 + \boxed{0.2} \\
&2^3 \times x = (100\ 1100)_2 + 0.4 \\
&2^4 \times x = (1001\ 1000)_2 + 0.8 \\
&2^5 \times x = (1\ 0011\ 0000)_2 + 1.6 = (1\ 0011\ 0001)_2 + 0.6 \\
&2^6 \times x = (10\ 0110\ 0010)_2 + 1.2 = (10\ 0110\ 0011)_2 + \boxed{0.2}
\end{aligned}$$

On a donc :

$$x = (1001, 10\ 0011)_2 + 2^{-6} \times 0.2$$

Si on continue la procédure ci-dessus, on remarque que pour $k \geqslant 2$ la partie fractionnaire de $2^k \times x$ est périodique : 0.2, 0.4, 0.8, puis 0.6. À chaque période on ajoute 0011 à droite de la partie entière de $2^k \times x$. En conclusion :

$$\boxed{9.55 = (1001, 10\ 0011\ 0011\ 0011\ldots)_2}$$