Exercice 1. Entiers signés

Question 1 – Ce sont les nombres compris entre -2^{63} et $2^{63}-1$ inclus.

Question 2 –

⋆ Pour $n = 80 \geqslant 0$. On a :

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

$$\boxed{\text{La représentation complément à deux sur 64 bits de l'entier 80 est donc } (\underbrace{0\ldots0}_{\text{57 zéros}}\ 101\ 0000).}$$

⋆ Pour $n = -100 < 0$. On a :

$$-n-1 = 99 = 64+32+2+1 = (110\ 0011)_2 = (\underbrace{0\ldots0}_{\text{57 zéros}}\ 110\ 0011)_2 \xrightarrow{flips} (\underbrace{1\ldots1}_{\text{57 uns}}\ 001\ 1100)_2$$

$$\boxed{\text{La représentation complément à deux sur 64 bits de l'entier } -100 \text{ est donc } (\underbrace{1\ldots1}_{\text{57 uns}}\ 001\ 1100)_2.}$$

Question 3 –

⋆ Pour $(\underbrace{0\ldots0}_{\text{57 zéros}}\ 100\ 1001)$. Soit n le nombre représenté en complément à deux. Comme le bit de poids fort est 0, on a $n \geqslant 0$ et :

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

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

⋆ Pour $(\underbrace{1\ldots1}_{\text{58 uns}}\ 00\ 1001)$. Soit n le nombre représenté en complément à deux. Comme le bit de poids fort est 1, on a $n < 0$ et :

$$(\underbrace{1\ldots1}_{\text{58 uns}}\ 00\ 1001) \xrightarrow{flips} (\underbrace{0\ldots0}_{\text{58 zéros}}\ 11\ 0110)$$

Donc :

$$-n-1 = (\underbrace{0\ldots0}_{\text{58 zéros}}\ 11\ 0110) = (11\ 0110) = 2+4+16+32 = 54$$

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

Exercice 2. Conversion de réels

Question 1 –

⋆ Pour -7.6875 :

$$-7.6875 = -(4+2+1+0.5+0.125+0.0625) = -(111.1011)_2 = (-1)^1(1.1110\,11)_2 \times 2^2$$

De plus, la représentation biaisée sur 11 bits de l'exposant $e = 2$ est la représentation binaire de $e + 2^{10} - 1 = 2^{10} + 1 = (1\underbrace{0\ldots0}_{\text{9 zéros}}1)_2$. Finalement :

$$\boxed{\text{La représentation } \texttt{binary64} \text{ de } -7.6875 \text{ est } (1 \mid 1\underbrace{0\ldots0}_{\text{9 zéros}}1 \mid 1110\,11\,\underbrace{0\ldots0}_{\text{46 zéros}})}$$

⋆ Pour -2^{-1022} :

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

De plus, la représentation biaisée sur 11 bits de l'exposant $e = -1022$ est la représentation binaire de $e + 2^{10} - 1 = 1 = (\underbrace{0\ldots0}_{\text{10 zéros}}\,1)_2$. Finalement :

$$\boxed{\text{La représentation } \texttt{binary64} \text{ de } 2^{-1022} \text{ est } (0 \mid \underbrace{0\ldots0}_{\text{10 zéros}}\,1 \mid \underbrace{0\ldots0}_{\text{52 zéros}})}$$

Question 2 –

⋆ Pour $(0 \mid 1\underbrace{0\ldots0}_{\text{8 zéros}}10 \mid 1101\,\underbrace{0\ldots0}_{\text{48 zéros}})$: l'exposant e dont la représentation biaisée sur 11 bits est $(1\underbrace{0\ldots0}_{\text{8 zéros}}10)$ vérifie :

$$e + 2^{10} - 1 = (1\underbrace{0\ldots0}_{\text{8 zéros}}10)_2.$$

Donc :

$$e = 1 + (10)_2 = 3.$$

Le nombre recherché est donc :

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

⋆ Pour $(1 \mid 0\underbrace{1\ldots1}_{\text{9 uns}}0 \mid 01\,\underbrace{0\ldots0}_{\text{50 zéros}})$: l'exposant e dont la représentation biaisée sur 11 bits est $(0\underbrace{1\ldots1}_{\text{9 uns}}0)$ vérifie :

$$e + 2^{10} - 1 = (0\underbrace{1\ldots1}_{\text{9 uns}}0)_2.$$

Donc :

$$-e = (1\,\underbrace{0\ldots0}_{\text{10 zéros}})_2 - (0\underbrace{1\ldots1}_{\text{9 uns}}0)_2 - 1 = (10)_2 - 1 = 1.$$

Donc $e = -1$. Le nombre recherché est donc :

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

Question 3 – 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 4 – 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 5 – 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 :

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