Représentation approximative des nombres réels

Corrigé

Exercice 1

Schéma de la conversion d'un nombre réel en IEEE 754 (simple précision)

1. **Exemple : Convertir 5.75 en IEEE 754**

Étape 1 : Représenter le nombre en binaire

1. Convertir 5 en binaire :

$$5_{10} = 101_2$$

2. Convertir 0.75 en binaire :

$$0.75 \times 2 = 1.5 \quad (\text{partie entière} = 1)$$
$$0.5 \times 2 = 1.0 \quad (\text{partie entière} = 1)$$

Donc, $0.75_{10} = 0.11_2$.

3. Donc, $5.75_{10} = 101.11_2$.

Étape 2 : Normaliser la représentation binaire

On écrit 101.11_2 sous forme normalisée (comme $1.M \times 2^E$) :

$$101.11_2 = 1.0111_2 \times 2^2$$

- La mantisse M est 1.0111_2 (on garde uniquement les chiffres après la virgule).
- L'exposant E est 2.

Étape 3 : Calculer l'exposant et le décalage

L'exposant E est représenté en binaire avec un **décalage de 127**. Donc on calcule :

$$E + 127 = 2 + 127 = 129$$

En binaire, $129_{10} = 10000001_2$.

Étape 4 : Remplir les 32 bits

- **Signe (1 bit)** : Le nombre est positif, donc le bit de signe est **0**.
- **Exposant (8 bits)** : $E + 127 = 10000001_2$ (8 bits).
- **Mantisse (23 bits)** : La mantisse est $01110000000000000000000_2$ (on garde les 23 premiers chiffres après la virgule).

Étapes de la conversion de $-9,42$ en IEEE 754

1. **Représentation en binaire**

Commençons par décomposer le nombre en partie entière et partie fractionnaire :

- $-9,42 = -9 - 0,42$

1.1. Partie entière : 9

Convertissons 9 en binaire :

$$9_{10} = 1001_2$$

1.2. Partie fractionnaire : $0,42$

Convertissons $0,42$ en binaire :

$$0,42 \times 2 = 0,84 \quad (\text{partie entière} = 0)$$
$$0,84 \times 2 = 1,68 \quad (\text{partie entière} = 1)$$
$$0,68 \times 2 = 1,36 \quad (\text{partie entière} = 1)$$
$$0,36 \times 2 = 0,72 \quad (\text{partie entière} = 0)$$
$$0,72 \times 2 = 1,44 \quad (\text{partie entière} = 1)$$
$$0,44 \times 2 = 0,88 \quad (\text{partie entière} = 0)$$

Et ainsi de suite... La représentation binaire de $0,42$ devient :

$$0,42_{10} \approx 0.0110100010001111_2$$

Et ainsi de suite... La représentation binaire de $0,42$ devient :

$$0,42_{10} \approx 0.0110100010001111_2$$

Donc, $9,42_{10} = 1001.0110100010001111_2$.

2. **Normalisation**

La représentation normalisée du nombre binaire est de la forme $1.M \times 2^E$.

$$1001.0110100010001111_2 = 1.0010110100010001111_2 \times 2^3$$

La mantisse M est 1.0010110100010001111_2 et l'exposant E est 3.

3. **Calcul de l'exposant (avec décalage de 127)**

Dans le format IEEE 754, l'exposant est stocké avec un **décalage de 127**.

$$E + 127 = 3 + 127 = 130$$

En binaire, $130_{10} = 10000010_2$.

4. **Gérer le signe**

Le nombre est **négatif**, donc le bit de signe $s = 1$.

5. **Remplir les 32 bits**

- **Signe** (1 bit) : Le nombre est négatif, donc $s = 1$.
- **Exposant** (8 bits) : $E + 127 = 130$, soit 10000010_2.
- **Mantisse** (23 bits) : On prend la partie après la virgule de la mantisse, soit 0010110100010001111_2 (complété avec des zéros à droite pour obtenir 23 bits).

Explication du résultat

- **Signe** : 1, car le nombre est négatif.
- **Exposant** : 10000010_2 (équivalent à 130 en décimal). Cela représente 2^3 après avoir appliqué le décalage de 127.
- **Mantisse** : La partie après la virgule du nombre normalisé, $00101101000100011110000_2$, est stockée en utilisant les 23 premiers bits.

Exercices 2 et 3

On a $128 = 1,0 \times 2^7$, donc $s = 0$, la mantise m vaut 0 et l'exposant (biaisé) $e = 127 + 7 = 134$, c'est-à-dire `10000110` en binaire. Au final, la représentation de 128 en simple précision est `0 10000110 00000000000000000000000`.

Puisque -32.75 est négatif, on a $s = 1$. Ensuite, 32 en binaire vaut `100000` et $0,75 = 1/2 + 1/2^2$, donc $32,75$ vaut en binaire $100000,11$, soit $1,0000011 \times 2^5$. La mantisse m vaut donc $000001100\ldots0$ et l'exposant biaisé $e = 127 + 5 = 132$. Au final, la représentation de $-32,75$ en simple précision est `1 10000100 00000110000000000000000`

Le nombre `1 01111110 11110000000000000000000` nous donne un signe négatif $s = 1$, un exposant biaisé $e = 126 - 127 = -1$ et une mantisse (en binaire) $1,1111 \times 2^{-1} = 0,11111$, c'est-à-dire $1/2 + 1/4 + 1/8 + 1/16 + 1/32 = 0,96875$. Soit au final, une valeur décimale de -0.96875.

Le nombre `0 10000011 11100000000000000000000` nous donne un signe positif $s = 0$, un exposant biaisé $e = 131 - 127 = 4$ et un mantisse (en binaire) $1,111 \times 2^4 = 11110$. Soit au final, un valeur décimale de 30.