Chapitre 1 : Arithmétique des Ordinateurs

Chemseddine Chohra

2020/2021

Table des matières

En Base 10

- $x = \pm a \times 10^e$.
- $1 \leq a < 10$.
- e un nombre entier.

En Base 10

- $x = \pm a \times 10^e$.
- $1 \leq a < 10$.
- e un nombre entier.

Exemple

- $29 = 2.9 \times 10^1$.
- $-0.0048 = -4.8 \times 10^{-3}$.

Pourquoi?

- Écrire facilement de très grandes/petites valeurs.
- Occuper moins de place.

Pourquoi?

- Écrire facilement de très grandes/petites valeurs.
- Occuper moins de place.

Exemple

- $13700000000 = 1.37 \times 10^{10}$.
- $0.0000000000000000000000000016726 = 1.6726 \times 10^{-27}$.

Généralisation en Base 2

- $x = \pm a \times 2^e$.
- $1 \leq a < 2$.
- e un nombre entier.

Généralisation en Base 2

- $x = \pm a \times 2^e$.
- $1 \leq a < 2$.
- e un nombre entier.

A noter

- La partie entière vaut toujours "1".
- $a = 1.F$.
- $x = \pm 1.F \times 2^b$.

Nous allons écrire les nombres suivants sous la notation scientifique en base 2 :

- A = 10
- B = 0.625
- C = -52.75

A = 10

- $10\ /\ 2$: résultat $= 5$, reste $= 0$
- $5\ /\ 2$: résultat $= 2$, reste $= 1$
- $2\ /\ 2$: résultat $= 1$, reste $= 0$
- $1\ /\ 2$: résultat $= 0$, reste $= 1$

A = 10

- 10 / 2 : résultat = 5, reste = 0
- 5 / 2 : résultat = 2, reste = 1
- 2 / 2 : résultat = 1, reste = 0
- 1 / 2 : résultat = 0, reste = 1

Conversion et Notation Scientifique

$$A = (1010)_2 = 1.01 \times 2^3$$

B = 0.625

- $0.625 \times 2 = 1.25$
- $0.25 \times 2 = 0.5$
- $0.5 \times 2 = 1.0$

$B = 0.625$

- $0.625 \times 2 = 1.25$
- $0.25 \times 2 = 0.5$
- $0.5 \times 2 = 1.0$

Conversion et Notation Scientifique

$$B = (0.101)_2 = 1.01 \times 2^{-1}$$

C = -52.75, Partie Entière

- 52 / 2 : résultat = 26, reste = 0
- 26 / 2 : résultat = 13, reste = 0
- 13 / 2 : résultat = 6, reste = 1
- 6 / 2 : résultat = 3, reste = 0
- 3 / 2 : résultat = 1, reste = 1
- 1 / 2 : résultat = 0, reste = 1

C = -52.75, Partie Fractionnaire

- $0.75 \times 2 = 1.5$
- $0.5 \times 2 = 1.0$

C = -52.75, Partie Fractionnaire

- $0.75 \times 2 = 1.5$
- $0.5 \times 2 = 1.0$

Conversion et Notation Scientifique

$$C = -(110100.11)_2 = -1.1010011 \times 2^5$$

Codage des Formats Binaires

- A partir de la notation scientifique en base 2.
- Coder le signe, la mantisse et l'exposant.
- Combien de bit utiliser pour chaque partie?
  - Des formats standards existent.
  - *Binary64* et *Binary32* les plus utilisés.

Nombres Flottants

Format *Binary32*

- Coder un nombre flottant sur 32 bits.
- Type *float* en C et *single* en Matlab.
- Les 32 bits sont divisés comme suit :
  - 1 bit de signe.
  - 8 bits d'exposant.
  - 23 bits de mantisse.

Nombres Flottants

Format *Binary64*

- Coder un nombre flottant sur 64 bits.
- Type *double* en C et Matlab.
- Les 64 bits sont divisés comme suit :
  - 1 bit de signe.
  - 11 bits d'exposant.
  - 52 bits de mantisse.

Revenons sur le Codage

- Codage du signe :
  - Signe positif : 0.
  - Signe négatif : 1.
- Codage de la mantisse :
  - Le code binaire de la partie fractionnaire.
  - La partie entière vaut toujours 1.
- Codage de l'exposant e :
  - Le code binaire de $e + \beta$.
  - β appelé biais d'exposant.
  - pour un exposant de taille x, nous avons :

$$\beta = 2^{x-1} - 1$$

$$\pm 1.F \times 2^{e}$$

$$\pm 1.F \times 2^{e}$$

signe	exposant	mantisse
0 if $+$, 1 if -	$(e+\beta)_2$	F

Format *Binary8*

- Format non-standard.
- Les 8 bits sont divisés comme suit :
  - 1 bit de signe.
  - 4 bits d'exposant ($\beta = 2^{4-1} - 1 = 7$).
  - 3 bits de mantisse.

Exemple 01 : Codage de $A = 10$

- $A = +1.01 \times 2^3$.
- Codage
  - Signe positif : code $= 0$
  - Exposant $= 3$: $\beta + 3 = 7 + 3 = 10 = (1010)_2$
  - Mantisse $= 1.01$: code $= 010$

Exemple 01 : Codage de A = 10

- $A = +1.01 \times 2^3$.
- Codage
  - Signe positif : code $= 0$
  - Exposant $= 3$: $\beta + 3 = 7 + 3 = 10 = (1010)_2$
  - Mantisse $= 1.01$: code $= 010$

Code de A $\Rightarrow$

signe	exposant	mantisse
0	1010	010

Exemple 02 : Codage de B = -0.625

- $B = -1.01 \times 2^{-1}$.
- Codage
  - Signe positif : code = 1
  - Exposant = 3 : $\beta - 1 = 6 = (0110)_2$
  - Mantisse = 1.01 : code = 010

Exemple 02 : Codage de B = -0.625

- $B = -1.01 \times 2^{-1}$.
- Codage
  - Signe positif : code = 1
  - Exposant = 3 : $\beta - 1 = 6 = (0110)_2$
  - Mantisse = 1.01 : code = 010

Code de B $\Rightarrow$

signe	exposant	mantisse
1	0110	010

Décodage d'un Nombre Flottant

Pour décoder un nombre flottant il faut :

- Séparer les bits de signe (s), d'exposant (E) et mantisse (F).
- Extraire la notation scientifique binaire avec la formule :

$$(-1)^s \times 1.F \times 2^{E-\beta}$$

- Finalement, convertir le résultat en décimal.

Nous allons décoder les deux nombres flottants suivants :

- 10101010
- 01010101

Nous allons décoder les deux nombres flottants suivants :

- 10101010
- 01010101

Les deux nombres sont codés sur le format *Binary8*.

Décodage de 10101010

- Décomposition : 10101010

Décodage de 10101010

- Décomposition : 10101010
- Notation scientifique en base 2 :

$$(-1)^{s} \times 1.F \times 2^{E-\beta} = (-1)^{1} \times 1.010 \times 2^{5-7}$$
$$= -1.010 \times 2^{-2}$$

Décodage de 10101010

- Décomposition : 10101010
- Notation scientifique en base 2 :

$$(-1)^s \times 1.F \times 2^{E-\beta} = (-1)^1 \times 1.010 \times 2^{5-7}$$
$$= -1.010 \times 2^{-2}$$

- Convertir en décimal :

$$-1.01 \times 2^{-2} = -(0.0101)_2 = -0.3125$$

Décodage de 10101010

- Décomposition : 10101010
- Notation scientifique en base 2 :

$$(-1)^{s} \times 1.F \times 2^{E-\beta} = (-1)^{1} \times 1.010 \times 2^{5-7}$$
$$= -1.010 \times 2^{-2}$$

- Convertir en décimal :

$$-1.01 \times 2^{-2} = -(0.0101)_2 = -0.3125$$

Poids	2^1	2^0		2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
Chiffre	$\cdots$	0	.	0	1	0	1	$\cdots$

Décodage de 01010101

- Décomposition : 01010101

Décodage de 01010101

- Décomposition : 01010101
- Notation scientifique en base 2 :

$$(-1)^s \times 1.F \times 2^{E-\beta} = (-1)^0 \times 1.101 \times 2^{10-7}$$
$$= +1.101 \times 2^3$$

Décodage de 01010101

- Décomposition : 01010101
- Notation scientifique en base 2 :

$$(-1)^{s} \times 1.F \times 2^{E-\beta} = (-1)^{0} \times 1.101 \times 2^{10-7}$$
$$= +1.101 \times 2^{3}$$

- Convertir en décimal :

$$1.101 \times 2^{3} = (1101)_2 = 13$$

L'arrondi

Arrondi au Plus Proche

- Si la partie fractionnaire de la mantisse ne tient pas sur le format.
- Le nombre n'est pas **exactement** représentable.
- Coder le nombre représentable le plus proche.
- Une erreur d'arrondi est introduite.

Nous allons coder les nombres suivants :

- $(45)_{10} = 1.01101 \times 2^5$
- $(110)_{10} = 1.10111 \times 2^6$
- $(19)_{10} = 1.0011 \times 2^4$

L'arrondi

Arrondi de 45 sur *Binary8*

$$(45)_{10} = 1.01101 \times 2^5$$

$$1.011 < 1.01101 < 1.100$$

Comment Choisir?

- 1.011 est la mantisse représentable la plus proche.
- $|1.01101 - 1.011| < |1.100 - 1.01101|$.
- $fl(45) = 1.011 \times 2^5$.

L'arrondi

Arrondi de 110 sur *Binary8*

$$(110)_{10} = 1.10111 \times 2^6$$

$$1.101 < 1.10111 < 1.110$$

Comment Choisir?

- 1.110 est la mantisse représentable la plus proche.
- $|1.10111 - 1.101| > |1.110 - 1.10111|$.
- $fl(110) = 1.110 \times 2^6$.

$$(19)_{10} = 1.0011 \times 2^4$$

$$1.001 < 1.0011 < 1.010$$

Comment Choisir?

- $|1.0011 - 1.001| = |1.010 - 1.0011|$.
- On choisit l'arrondi avec une code de mantisse pair (0 à la fin).
- $fl(19) = 1.010 \times 2^4$.

Opérations Flottantes

Comment Choisir?

- Entrées : nombres flottants. Sortie : nombres flottants.
- L'arrondi de l'opération réelle correspondante.
- Opérations flottantes de base

$$a \oplus b = fl(a+b)$$
$$a \ominus b = fl(a-b)$$
$$a \otimes b = fl(a \times b)$$
$$a \oslash b = fl(a/b)$$

Addition et Soustraction

Étapes

- Aligner les mantisses (avoir le même exposant).
- Additionner les deux mantisses.
- Normaliser le résultat.
- Arrondir le résultat.

Calculer l'addition suivante :

$$\begin{array}{lcl} & 1.110 & \times 2^3 \\ + & 1.011 & \times 2^2 \\ \hline = & & \end{array}$$

Calculer l'addition suivante :

$$
\begin{array}{rll}
 & 1.110 & \times 2^3 \\
+ & 0.1011 & \times 2^3 \\
\hline
= & &
\end{array}
$$

Étapes

- Aligner les mantisses.

Calculer l'addition suivante :

$$
\begin{array}{lll}
 & 1.110 & \times 2^3 \\
+ & 0.1011 & \times 2^3 \\
\hline
= & 10.0111 & \times 2^3
\end{array}
$$

Étapes

- Additionner les mantisses.

Calculer l'addition suivante :

$$\begin{array}{lll} & 1.110 & \times 2^3 \\ + & 0.1011 & \times 2^3 \\ \hline = & 1.00111 & \times 2^4 \end{array}$$

Étapes

- Normalisation

Calculer l'addition suivante :

$$
\begin{array}{rll}
 & 1.110 & \times 2^3 \\
+ & 0.1011 & \times 2^3 \\
\hline
= & 1.010 & \times 2^4
\end{array}
$$

Étapes

- Arrondi

Multiplication et Division

$$M_1 \times 2^{E_1} \times M_2 \times 2^{E_2} = (M_1 \times M_2) \times 2^{E_1+E_2}$$
$$(M_1 \times 2^{E_1})/(M_2 \times 2^{E_2}) = (M_1/M_2) \times 2^{E_1-E_2}$$

Multiplication et Division

$$M_1 \times 2^{E_1} \times M_2 \times 2^{E_2} = (M_1 \times M_2) \times 2^{E_1+E_2}$$
$$(M_1 \times 2^{E_1})/(M_2 \times 2^{E_2}) = (M_1/M_2) \times 2^{E_1-E_2}$$

Étapes

- Additionner/Soustraire les exposants.
- Multiplier/Diviser les mantisses.
- Normaliser le résultat.
- Arrondir le résultat.

Calculer la multiplication suivante :

$$\begin{array}{lrl} & 1.110 & \times 2^3 \\ \times & 1.011 & \times 2^2 \\ \hline = & & \end{array}$$

Calculer la multiplication suivante :

$$
\begin{array}{cll}
 & 1.110 & \times 2^3 \\
\times & 1.011 & \times 2^2 \\
\hline
= & 10.011010 & \times 2^5
\end{array}
$$

Étapes

- Additionner les exposants et multiplier les mantisses

Calculer la multiplication suivante :

$$\begin{array}{ccc} & 1.110 & \times 2^3 \\ \times & 1.011 & \times 2^2 \\ \hline = & 1.0011010 & \times 2^6 \end{array}$$

Étapes

- Normalisation

Calculer la multiplication suivante :

$$\begin{array}{rrl} & 1.110 & \times 2^3 \\ \times & 1.011 & \times 2^2 \\ \hline = & 1.010 & \times 2^6 \end{array}$$

Étapes

- Arrondi

FIN!

Merci de Votre Attention