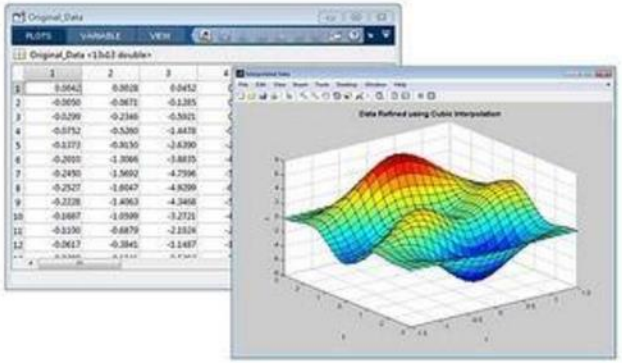
Méthodes Numériques

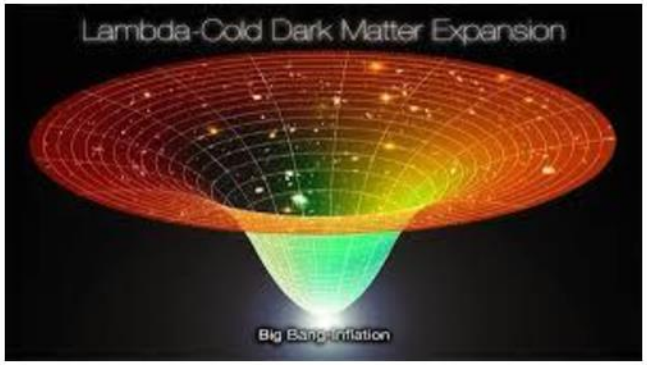
2iéme Année Lc. Inf.

Introduction

Motivation

- Calcul





- Bug

Motivation

Domaines :

- aéronautique,
- géométrie 3D,
- BTP, bâtiment et des travaux publics
- prévisions météo,
- simulations nucléaires,
- feuilles Excel,

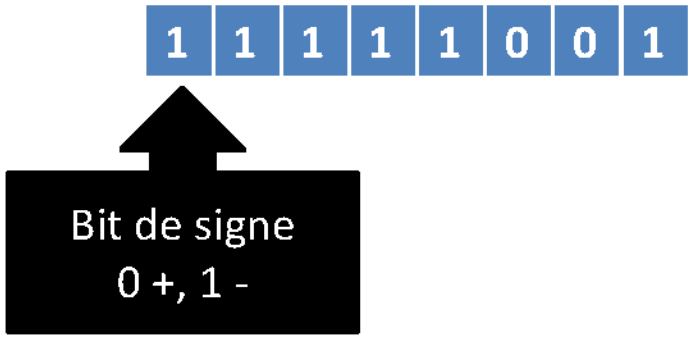
Représentation des nombre en machine

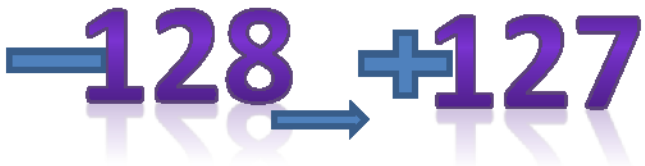
- Entiers Positifs / Négatifs
- 7, en décimal.
- 111, en binaire.
- 00000111, sur un mot mémoire de 8 bits.
- 0000000000000111, sur un mot mémoire de 16 bits.

Représentation des nombre en machine

- Entiers Positifs / Négatifs
- -7, en décimal.
- 11111001, sur un mot mémoire de 8 bits.

Méthode:
00000111, 7 en binaire
11111000 , inverser les chiffres (complément à 1)
11111001, +1 (complément à 2)





Représentation des nombre en machine

- Entiers Positifs / Négatifs
- -7, en décimal.
- 1111111111111001, sur un mot mémoire de 16 bits.

Méthode:
0000000000000111, 7 en binaire
1111111111111000 , inverser les chiffres (complément à 1)
1111111111111001, +1 (complément à 2)

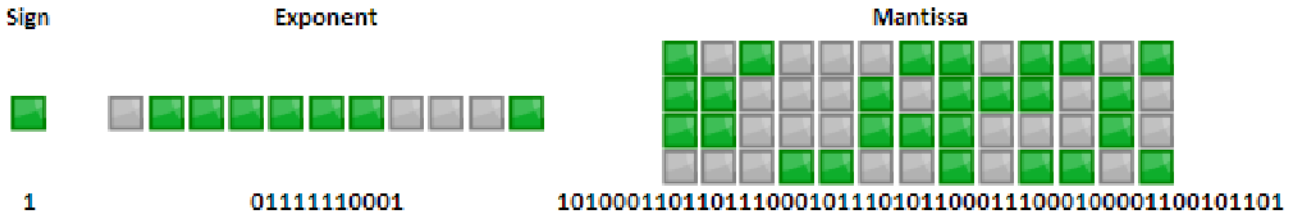
Représentation des nombre en machine

FPU

- Flottant?
- Réel: 0,5 0,0001 3,07
- Binaire (Norme IEEE 754) :
- -0,0001, décimal
- **10111000110100011011011100010111**

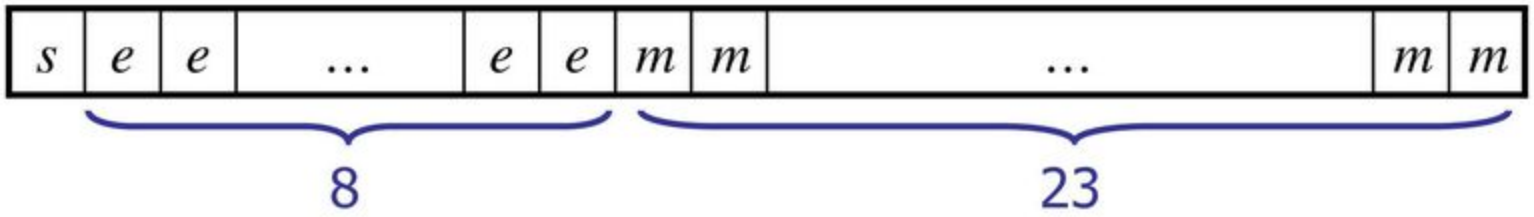
Représentation des nombre en machine

- **10111111000110100011011011100010111010110001110001000011001011 01**

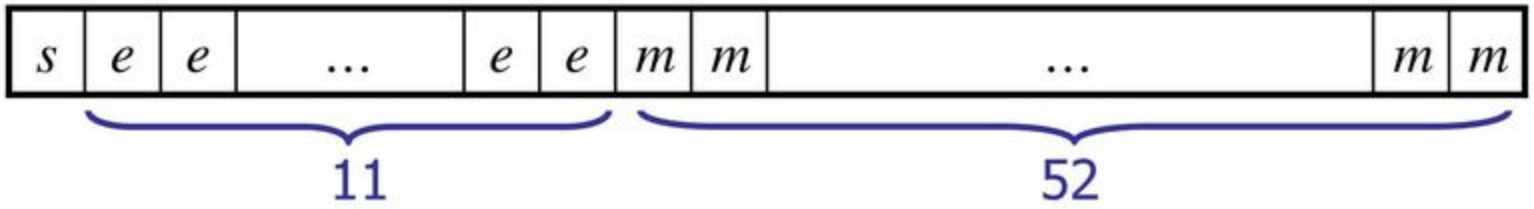


Le type REAL : Limite de représentation

- Réel simple précision
  - Motif binaire de la forme :



  - Nombre représenté : $r = s1.m\ 2^{e-127}$
  - 32 bits (4 octets): $1.2\ 10^{-38} \leq |x| \leq 3.4\ 10^{38}$
  - Précision : 7 chiffres significatifs
- Réel double précision
  - Motif binaire de la forme:



  - Nombre représenté : $r = s1.m\ 2^{e-1023}$
  - 64 bits (8 octets): $2.2\ 10^{-308} \leq |x| \leq 1.8\ 10^{308}$

Exemple

- 63,3 en virgule flottante simple précision

Diviseur	Nombre	Reste
2	263	-
2	131	1
2	65	1
2	32	1
2	16	0
2	8	0
2	4	0
2	2	0
2	1	0
2	0	1

0,3x2	0,6	0
0,6x2	1,2	1
0,2x2	0,4	0
0,4x2	0,8	0
0,8x2	1,6	1
0,6x2	1,2	1
……		

$63_{10} = 100000111_2$

$0{,}3_{10} = 010011\ldots_2$

$63{,}3_{10} = 100000111{,}010011\ldots_2$

$63{,}3_{10} = 1{,}00000111010011\ldots 2^8{}_2$

Exemple

- 63,3 en virgule flottante simple précision

  $63{,}3_{10} = 1{,}00000111010011\ldots 2^{8}{}_{2}$

- S=0
- M=00000111010011001100110
- E=127+8=135=10000111

  $63{,}3_{10} = 0 \quad 10000111 \quad 00000111010011001100110_{2}$

Exemple

- 63,3 en virgule flottante double précision

  $63{,}3_{10} = 1{,}00000111010011 \ldots 2^{8}{}_{2}$

- S=0
- M=0000011101001100110011001100110011001100110011001100
- E=1023+8=1023=10000000111

  $63{,}3_{10} = 0 \;\; 10000000111 \;\; 0000011101001100110011001100110011001100110011001100_{2}$

Représentation des nombre en machine

Exceptions :

- En arithmétique flottante IEEE, un calcul peut aboutir à des valeurs qui ne correspondent pas à des nombres réels :
- NaN (« not a number »), qui sera par exemple le résultat de la tentative de division flottante de zéro par zéro, ou de la racine carrée d'un nombre strictement négatif. Les NaN se propagent : la plupart des opérations faisant intervenir un NaN donnent NaN (des exceptions sont possibles, comme NaN puissance 0, qui peut donner 1).
- Un infini positif et un infini négatif, qui sont par exemple le résultat d'un débordement en arrondi au plus près.

Les erreurs de calculs

Erreur absolue et erreur relative

- Quantité exacte: 1, ¼, $\sqrt{2}$, π, log2, e, …
- Quantité approximative ou valeur approchée: $\sqrt{2}$=1.41, π=3.14, e=2.718, …
- Soient x une quantité exacte et x* une valeur approchée de x:
- Si x*> x alors x* est dite approchée par excès.
- Si x*< x alors x* est dite approchée par défaut.

Erreur absolue et erreur relative

- Exemple
- Considérons le nombre $\sqrt{2}$:
- 1.41 est une valeur approchée par défaut.
- 1.42 est une valeur approchée par excès.
- Nous avons l'encadrement suivant:
- $1.41 < \sqrt{2} < 1.42$.

Erreur absolue

- On appelle erreur absolue de x* (sur x), la quantité E=|x-x*|.
- Plus l'erreur absolue sur x* est petite, plus x* est précise.
- Exemple: x=2/3, la valeur approchée $x_1{}^*$=0.666667 est mille fois plus précise que la valeur approchée $x_2{}^*$=0,667.

Erreur absolue

- En effet:
- $E_1 = \left| x - {x_1}^* \right| = \left| \frac{2}{3} - 0.666667 \right| = \left| \frac{2000000 - 2000001}{3.10^6} \right| = \frac{1}{3}.10^{-6}$
- $E_2 = \left| x - {x_2}^* \right| = \left| \frac{2}{3} - 0.667 \right| = \left| \frac{2000 - 2001}{3.10^3} \right| = \frac{1}{3}.10^{-3}$
- $E_2 = E_1.10^3$ E_2 est 1000 fois plus grande que E_1
- Donc, ${x_1}^*$ est 1000 fois plus précise que ${x_2}^*$.

Erreur relative

- On appelle erreur relative à x* la quantité $E_r = \frac{|x - x^*|}{|x|} = \frac{E}{|x|}$
- E_r est souvent exprimée en pourcentage.
- Exemple:
- x=2/3 x*=0.67
- y=1/15 y*=0.07

Erreur relative

- $E_1 = |x - x_1^*| = \left|\frac{2}{3} - 0.67\right| = \left|\frac{200-201}{3.10^2}\right| = \frac{1}{3}.10^{-2}$
- $E_2 = |x - x_2^*| = \left|\frac{1}{15} - 0.07\right| = \left|\frac{100-105}{15.10^2}\right| = \frac{1}{3}.10^{-2}$
- $E_{r1} = \frac{E_1}{|x|} = 0.5\ 10^{-2} = 0{,}5\%$
- $E_{r2} = \frac{E_2}{|y|} = 5\ 10^{-2} = 5\%$
- Ainsi, bien que les erreurs relatives soient égales, x* est une approximation 10 fois plus précise pour x que y* l'est pour y.

Majorants des erreurs absolue et relative

On appelle majorant de l'erreur absolue dune valeur approchée x^* tout nombre réel positif $\triangle x$ vérifiant :
$E = |x - x^*| \leq \triangle x$ ou de manière équivalente : $x^* - \triangle x \leq x \leq x^* + \triangle x$. On écrit $x = x^* \pm \triangle x$

$\partial x = \frac{\Delta x}{|x^*|}$, ∂x est appelé à défaut l'erreur relative de x* et est exprimé souvent en %

Propagation des erreurs

Soient x et y deux valeur exactes, x^* et y^* deux approximations de x et y, $\triangle x$ et $\triangle y$ les erreurs absolues et δx et δy les erreurs relatives.

Addition

$$\triangle(x+y) = \triangle x + \triangle y \text{ et } \delta(x+y) \leq max(\delta x, \delta y)$$

Soustraction

$$\triangle(x-y) = \triangle x + \triangle y \text{ et } \delta(x-y) \leq \frac{|x^*+y^*|}{|x^*-y^*|} max(\delta x, \delta y)$$

Multiplication

$$\triangle(xy) = x^* \triangle y + y^* \triangle x \text{ et } \delta(xy) = \delta x + \delta y$$

Division

$$\triangle(x/y) = \frac{x^* \triangle y + y^* \triangle x}{(y^*)^2} \text{ et } \delta(x/y) = \delta x + \delta y$$

- Round
- Truncate
- Overflow
- Underflow