

Representación de números en una computadora

- Sistemas de numeración posicional: Fijo y flotante.
- Número de máquina y rango de máquina.
- Unidad de redondeo.
- Errores de redondeo en operaciones aritméticas en punto flotante.
- Representación de números en la computadora en sistema flotante.

Sistemas de numeración

Sistema posicional: El valor depende de la posición absoluta de cada uno de los símbolos que lo contienen.

Los caracteriza una base B (por ejemplo 2, 10, 16, etc)

Conjunto elementos de la base: $0, 1, \dots, B-1$

La convención para expresar un número x en base 10 por ejemplo sería el siguiente:

$$x = \underbrace{\pm}_{\text{signo}} \underbrace{a_M a_{M-1} \dots a_0}_{\text{Parte entera}}, \underbrace{a_{-1} a_{-2} \dots a_{-N}}_{\text{Parte decimal}} \quad a_k \text{ es un número entre 0 y 9}$$

Estamos simbolizando:

$$x = \pm \sum_{k=-N}^M a_k 10^k$$

Ejemplo de representación en notación decimal y binaria

En base 10

$$3.25 = + (3 \times 10^0) + (2 \times 10^{-1}) + (5 \times 10^{-2})$$

$$X = + (a_0 \times 10^0) + (a_{-1} \times 10^{-1}) + (a_{-2} \times 10^{-2})$$

$$3.25 = 3.25_{\text{diez}}$$

En base 2

$$3.25 = + (1 \times 2^1) + (1 \times 2^0) + (0 \times 2^{-1}) + (1 \times 2^{-2})$$

$$3.25 = 11.01_{\text{dos}}$$

Representación en sistema de numeración de punto fijo

Se toman dos números fijos n_1 y n_2

n_1 : lugares a los dígitos enteros.

n_2 : lugares a los dígitos decimales.

$n = n_1 + n_2$: Cantidad de lugares totales.

Si $n = 10$, $n_1 = 4$, y $n_2 = 6$:

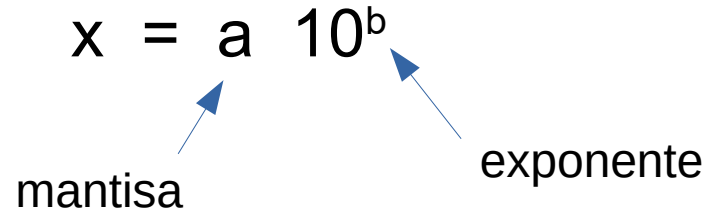
3.1415 se representará 0003 141500

1.0736 se representará 0001 073600

El número 123456 no lo podemos representar.

(Usualmente $n=n_1$ $n_2 = 0$)

Representación en sistema de numeración de punto flotante

$$x = a \cdot 10^b$$


mantisa

exponente

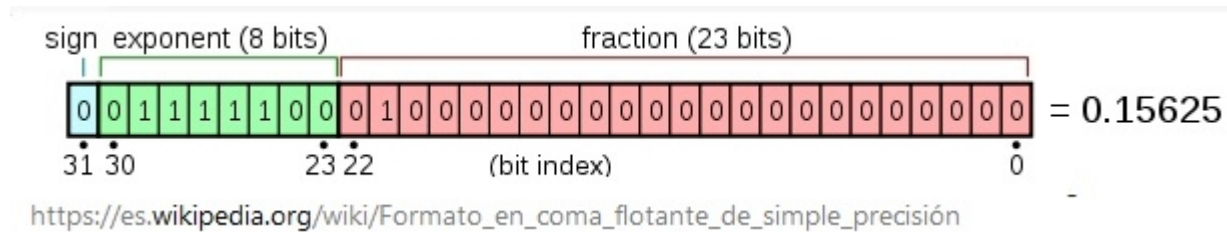
Si pedimos: $0.1 \leq |a| < 1$ Se denomina "punto flotante normalizado"

Para un sistema de base B . $\frac{1}{|B|} \leq |a| < 1$

Una computadora asignaría una cantidad fija y finita de t cifras para la mantisa y otra de e cifras para el exponente

Simple precisión en una computadora

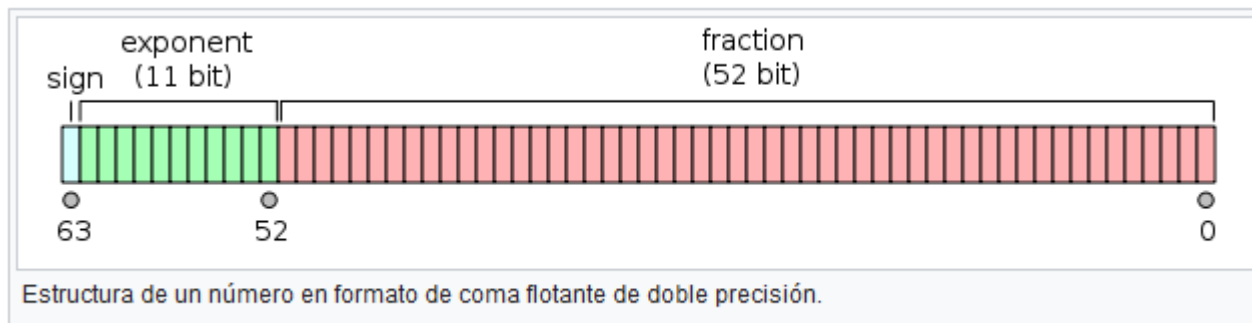
El formato en coma flotante de simple precisión es un formato que ocupa 32 bits (4 bytes) en su memoria



1 bit para el signo: $2^1=2$ representa + y -
8 bits para el exponente desplazado: $2^8=256$ se desplaza en 127
23 bits para la mantisa: 2^{23} se pueden representar $2^{24} = 16.777.216 = 10^{7.2}$

Doble precisión en una computadora

El formato en coma flotante de doble precisión es un formato que ocupa 64 bits (8 bytes) en su memoria



https://es.wikipedia.org/wiki/Formato_en_coma_flotante_de_doble_precisi3n

1 bit para el signo: $2^1=2$ representa + y -
11 bits para el exponente desplazado: $2^{11}=2048$ se desplaza 1023
52 bits para la mantisa: 2^{52} se pueden representar $2^{53} = 10^{15.9}$

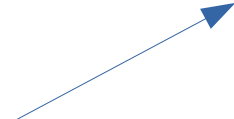
Número de máquina y rango de máquina

Números de máquina: Es el conjunto de números que se pueden representar exactamente. Esto quedará determinado por los números t y e , junto a la base B .

Rango de la máquina: Es el conjunto de números que podemos representar aproximadamente en nuestra máquina.

$$Rango = \{ |x| \in \mathbb{R} / x \in [10^{-10^e}, 10^{10^e-1}] \}$$

Errores inherentes o de representación en la computadora

$$\frac{1}{10} = 0.0001100110011001100110011 \dots_{dos}$$


No es número de máquina trabajando en simple o doble precisión

Si queremos realizar $z = x - y$

$x = 3.1416$ no es número de máquina

$y = 3.1415$ no es número de máquina

Hay dos errores: inherentes en el ingreso de los valores y el error en la operación

Errores de redondeo en operaciones aritméticas en punto flotante.

El error relativo de las operaciones elementales se pueden acotar por μ