

Punto flotante: introducción

Requisitos. Notación científica, sistema binario y hexadecimal.

Representación de números en punto flotante en computadoras

Ligas útiles:

http://www.fdi.ucm.es/profesor/mozos/AEC/aritm_pf.PDF

<http://babbage.cs.qc.edu/IEEE-754/>

Precisión doble. Usa 64 bits: 1 bit de signo, 11 de exponente, 52 de mantisa.

$$(-1)^{S'} \cdot 2^{E'-1023} \cdot 1.M'$$

1. Calcular el número cuya representación en punto flotante de precisión doble es

```
0 10000000100 011011000000000000000000000000000000000000000000000000000000
```

2. **Tarea.** Calcular el número cuya representación en punto flotante de precisión doble es

```
0 10000000011 100100000000000000000000000000000000000000000000000000000000
```

3. Escribir la secuencia de bits que representa el número 11.75 en punto flotante de precisión doble.

4. **Tarea.** Escribir la secuencia de bits que representa el número 0.1875 en punto flotante de precisión doble.

Algunas codificaciones con significado especial.

- Para +0, $S = 0$, E' consiste en ceros y M' consiste en ceros.
- Para $+\infty$, $S = 0$, E' consiste en unos y M' consiste en ceros.
- Para +NaN (Positive Not a Number), $S = 0$, E' consiste en unos y M' no consiste solamente en ceros.

Redondeo. El estándar IEEE-754 exige que el resultado de las operaciones sea el mismo que se obtendría si se realizasen con precisión absoluta y después se redondease. Hay 4 modos de redondeo:

- redondeo a cero (*truncamiento*);
- redondeo al más cercano (al par en caso de empate), llamado brevemente *redondeo*;
- redondeo a más infinito (por exceso);
- redondeo a menos infinito (por defecto).

5. Redondear los siguientes números binarios a 3 dígitos después del punto, usando varios modos de redondeo:

1.010001 1.01111 1.0001 1.0011