

¿Qué es el sistema IEEE 754?

El sistema IEEE 754 es un estándar para la representación de números en punto flotante. Este sistema establece las características de las tres partes de un número en punto flotante: el **signo**, **exponente** y **fracción (mantisa)**. Se enfoca en el método de precisión simple que utiliza 32 bits.

Características

Bit de signo:

El bit de signo es el primer bit del número en punto flotante. **Si el bit de signo es 0**, el número es **positivo**, **si es 1**, el número es **negativo**.

Exponente:

Indica cuántos “lugares” se debe desplazar hacia la derecha o hacia la izquierda la coma binaria de la parte significativa. El exponente de un número puede ser tanto positivo como negativo. En el caso de la precisión simple, el exponente se representa con 8 bits y se suma 127 al valor del exponente para obtener el valor.

Mantisa o parte fraccionaria:

La mantisa es la parte fraccionaria del número en punto flotante. En el caso de la precisión simple, la mantisa se representa con 23 bits. Debe ser normalizada, eso se logra moviendo la coma binaria a la izquierda hasta que el primer bit sea 1.

$$1,xxxx \times 2^{yyyy}$$

Al hacer esto, se debe adaptar el exponente para que el valor no se modifique, es decir, el exponente será ahora **la cantidad de veces que se movió la coma binaria a la izquierda**.

Convertir de decimal a IEEE 754

- Determinar el bit de signo.
- Convertir la parte entera y la parte fraccionaria a binario.
- Normalizar.
- Encontrar el exponente moviendo la coma y sumando 127, expresado en binario.
- Encontrar la parte fraccionaria.
- Conformar el número en punto flotante.

Convertir de IEEE 754 a decimal

- Dividir el número en punto flotante en sus tres partes.
- Encontrar el bit de signo.
- Encontrar el exponente y restarle 127.
- Desnormalizar el número y pasarlo a decimal.

Ejemplo de conversión de decimal a IEEE 754

Se busca convertir el número 723.125_{10} a IEEE 754.

- El bit de signo es 0, ya que el número es positivo.
- Se convierte la parte entera y la parte fraccionaria a binario.

- Parte entera:

$$\begin{aligned}723 \div 2 &= 361 \text{ residuo } 1 \\361 \div 2 &= 180 \text{ residuo } 1 \\180 \div 2 &= 90 \text{ residuo } 0 \\90 \div 2 &= 45 \text{ residuo } 0 \\45 \div 2 &= 22 \text{ residuo } 1 \\22 \div 2 &= 11 \text{ residuo } 0 \\11 \div 2 &= 5 \text{ residuo } 1 \\5 \div 2 &= 2 \text{ residuo } 1 \\2 \div 2 &= 1 \text{ residuo } 0 \\1 \div 2 &= 0 \text{ residuo } 1\end{aligned}$$

Por lo que la parte entera en binario es 1011010011.

- Parte fraccionaria:

$$\begin{aligned}0.125 \times 2 &= 0.25 \\0.25 \times 2 &= 0.5 \\0.5 \times 2 &= 1\end{aligned}$$

Por lo que la parte fraccionaria en binario es 001.

El número en binario es 1011010011.001.

- Para normalizar movemos la coma tantos lugares como sea necesario para que quede un uno seguido por una parte fraccionaria.

- Sin normalizar:** $1011010011.001 \times 2^0$
- Normalizado:** $1.011010011001 \times 2^{10}$

- El exponente es la cantidad de veces que se movió la coma binaria a la izquierda, en este caso, 10 veces. Por lo que el exponente es $127 + 10 = 137_{10} = 10001001_2$.

- La parte fraccionaria se compone por los 23 bits que siguen a la coma decimal del número normalizado.

$$01101001100100000000000$$

- Para conformar el número en punto flotante, se coloca el bit de signo, el exponente y la parte fraccionaria.

$$0 \ 10001001 \ 01101001100100000000000$$