



ESCUELA POLITÉCNICA NACIONAL
FACULTAD DE INGENIERÍA DE SISTEMAS
INGENIERÍA EN CIENCIAS DE LA COMPUTACIÓN

PERÍODO ACADÉMICO: 2025-A

ASIGNATURA: ICCD412 Métodos Numéricos

GRUPO: GR2

TIPO DE INSTRUMENTO: Práctica2

FECHA DE ENTREGA LÍMITE: 04/05/2025

ALUMNO: Sebastián Chicaiza

TEMA

Representación numérica 64 bits

OBJETIVOS

- Comprender el procedimiento para la conversión de números en decimal a formato IEEE 754 de 64 bits.
- Analizar el error relativo resultante de la conversión.

MARCO TEÓRICO

El estándar IEEE 754 define cómo representar números reales en computadoras usando formatos de punto flotante. Los tipos más comunes son float y double. El valor cero no puede representarse en formato normalizado, por lo que se usa un patrón especial de bits en ceros. También existe el -0, pero tanto 0 como -0 se consideran iguales al compararlos [1].

El número representado en formato IEEE 754 de 64 bits esá dado por:

$$x = (-1)^{Signo} \cdot (1 + Mantisa) \cdot 2^{exponente-1023}$$

[2]

DESARROLLO

Pasar el número $263,3_{10}$ al formato IEEE 754 de 64 bits, una vez en binario pasarlo a decimal y calcular el error relativo a 3 cifras de redondeo.

Transformación a IEE 754

Signo positivo (0)

Conversión a binario:

$$\frac{263}{2} = residuo(1)$$

$$\frac{131}{2} = residuo(1)$$

$$\frac{65}{2} = residuo(1)$$

$$\frac{32}{2} = residuo(0)$$

$$\frac{16}{2} = residuo(0)$$

$$\frac{8}{2} = residuo(0)$$

$$\frac{4}{2} = residuo(0)$$

$$\frac{2}{2} = residuo(0)$$

$$\frac{1}{2} = residuo(1)$$

$$263_{10} = 100000111_2$$

$$0,3 \cdot 2 = parteEntera(0)$$

$$0,6 \cdot 2 = parteEntera(1)$$

$$0,2 \cdot 2 = parteEntera(0)$$

$$0,4 \cdot 2 = parteEntera(0)$$

$$0,8 \cdot 2 = parteEntera(1)$$

$$0,6 \cdot 2 = parteEntera(1) \quad Se \quad repite$$

$$0,3_{10} = 0,01001_2$$

Transformación a decimal

$$\begin{aligned}x &= (-1)^{Signo} \cdot (1 + Mantisa) \cdot 2^{exponente-1023} \\x &= (-1)^0 \cdot (1 + 0,028515338897705078) \cdot 2^{1031-1023} \\x &= 263,2999267578125\end{aligned}$$

Cálculo de error relativo

$$\begin{aligned}p &= 263,3 \quad \text{y} \quad p^* = 263,2999267578125 \\error_{relativo} &= \left| \frac{p - p^*}{p} \right| \\&= \left| \frac{263,3 - 263,2999267578125}{263,3} \right| \\&= 0,000000278170100688829 \approx 0,278 \cdot 10^{-6}\end{aligned}$$

CONCLUSIONES

- La conversión a formato IEEE 754 de 64 bits permite representar una gran cantidad de valores con alta precisión.
- El resultado del cálculo del error relativo de la conversión demuestra la precisión de este formato.

RECOMENDACIONES

Tener en cuenta el valor del error relativo al utilizar el formato IEEE 754 para representar números. Especialmente cuando se trata de resultados en el campo científico, donde se requiere de gran precisión.

REFERENCIAS

- [1] Microsoft. (2023) Ieee 754 floating-point representation. Accessed: 2025-05-04. [Online]. Available: <https://learn.microsoft.com/en-us/dotnet/standard/numerics/ieee-754-floating-point>
- [2] H. Bui and S. Tahar, “Design and synthesis of an ieee-754 exponential function,” in *Engineering Solutions for the Next Millennium. 1999 IEEE Canadian Conference on Electrical and Computer Engineering (Cat. No. 99TH8411)*, vol. 1. IEEE, 1999, pp. 450–455.