

Representación flotante

Floating representation

Oscar Julián Cárdenas Salgado
Universidad tecnológica de Pereira
Julian.cardenas1@utp.edu.co

Resumen— En este documento hablaremos a cerca de la representación flotante o coma flotante, de sus funciones y de cómo podemos agrupar una cantidad grande de números en un número reducido de casillas.

Palabras clave— Coeficiente, base, exponente, signo, binario, notación científica.

Abstract— In this document we will talk about the floating representation or floating point, its functions and how we can group a large number of numbers in a small number of boxes.

Key Word — Coefficient, base, exponent, sign, binary, scientific notation.

I. INTRODUCCIÓN

Los computadores no pueden representar los números reales $\{R\}$ o los números complejos $\{C\}$ de manera exacta. Los computadores aproximan los $\{R\}$ mediante el sistema de números de punto flotante, utilizando un número finito de bits para su representación. [1]

Estos números se pueden representar por medio de notación científica, ya que esta notación permite que los computadores puedan representar números reales, sin importar si estos son muy grandes o muy pequeños, la notación científica permite que estos números se puedan agrupar de manera más eficiente y compacta, permitiéndole de esta manera a la computadora realizar operaciones aritméticas

Muchas aplicaciones requieren de números que se mueven en rangos muy grandes, por lo que en estos casos utilizamos la notación científica con el número flotante, para así comprimir y agrupar estos números y de esta manera el computador pueda leerlos.

II. CONTENIDO

1) Estándar. Casi todo el hardware y lenguajes de programación utilizan números de punto flotante en los mismos formatos binarios, que están definidos en el estándar IEEE 754. Los formatos más comunes son de 32 o 64 bits de longitud total. [2]

El estándar IEEE 754 ha sido definido por el Instituto de Ingenieros Eléctricos y Electrónicos (Institute of Electrical and Electronics Engineers, IEEE) y establece dos formatos básicos para representar a los números reales en la computadora digital: precisión simple y precisión doble. [3]

2) ¿Cómo funcionan los números del punto flotante? La idea es descomponer el número en dos partes:

- Una mantisa (también llamada coeficiente o significando) que contiene los dígitos del número. Mantisas negativas representan números negativos.
- Un exponente que indica dónde se coloca el punto decimal (o binario) en relación al inicio de la mantisa. Exponentes negativos representan números menores que uno.

Este formato cumple todos los requisitos:

- Puede representar números de órdenes de magnitud enormemente dispares (limitado por la longitud del exponente).
- Proporciona la misma precisión relativa para todos los órdenes (limitado por la longitud de la mantisa).
- Permite cálculos entre magnitudes: multiplicar un número muy grande y uno muy pequeño conserva la precisión de ambos en el resultado. [5]

3) Algunos datos del punto flotante:

- La secuencia de bits es primero el bit del signo, seguido del exponente y finalmente los bits significativos.
- El exponente no tiene signo; en su lugar se le resta un desplazamiento (127 para sencilla y 1023 para doble precisión). Esto, junto con la secuencia de bits, permite que los números de punto flotante se puedan comparar y ordenar correctamente incluso cuando se interpretan como enteros.
- Se asume que el bit más significativo de la mantisa es 1 y se omite, excepto para casos especiales.
- Hay valores diferentes para cero positivo y cero negativo. Estos difieren en el bit del signo, mientras

que todos los demás son 0. Deben ser considerados iguales aunque sus secuencias de bits sean diferentes.

- Hay valores especiales no numéricos (NaN, «not a number» en inglés) en los que el exponente es todo unos y la mantisa no es todo ceros. Estos valores representan el resultado de algunas operaciones indefinidas (como multiplicar 0 por infinito, operaciones que involucren NaN, o casos específicos). Incluso valores NaN con idéntica secuencia de bits no deben ser considerados iguales. [7]

4) Antes de entrar con detalle en la aritmética de los números en punto flotante, es interesante notar una propiedad de estos números de especial importancia en los cálculos numéricos y que hace referencia a su densidad en la línea real. Supongamos que p , el número de bits de la mantisa, sea 24. En el intervalo $\left[\frac{1}{2}, 1\right)$ (exponente $f = 0$) es posible representar 224 números igualmente espaciados y separados por una distancia $1/224$. De modo análogo, en cualquier intervalo $\left[2^f, 2^{f+1}\right)$ hay 224 números equiespaciados, pero su densidad en este caso es $2^f/224$. Por ejemplo, entre $220 = 1048576$ y $221 = 2097152$ hay $224 = 16777216$ números, pero el espaciado entre dos números sucesivos es de sólo $\frac{1}{16}$. De este hecho se deriva inmediatamente una regla práctica: cuando es necesario comparar dos números en punto flotante relativamente grandes, es siempre preferible comparar la diferencia relativa a la magnitud de los números. En la figura (1) se representa gráficamente la separación entre dos números consecutivos en función del exponente f en el rango $f = [20, 30]$. [8]

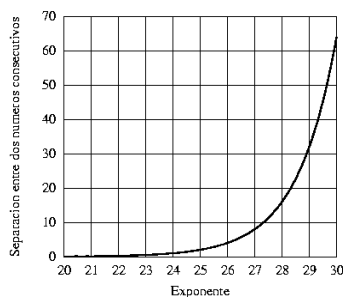
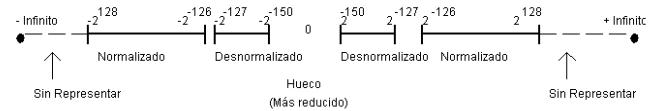


Figure: Evolución de la separación entre dos números consecutivos en función del exponente, f , de la representación en punto flotante de un número real.
[bb=55 60 455 410, clip=true, scale=0.7]eps/expon

III. CONCLUSIONE

Como conclusión expresamos que la representación en coma flotante se utiliza para representar a los números reales. Sin embargo no puede representar a todos ya que estos son infinitos y la capacidad de almacenamiento en coma flotante es finita. El exponente determina el rango de la recta real, es

decir los segmentos de ella, mientras que la mantisa determina la precisión (longitud del mínimo intervalo). [9]



Representación de la recta real en el IEEE 754

REFERENCIAS

- [1] <https://www.inf.utfsm.cl/~parce/cc1/clase18-RP.html>
- [3] <https://medium.com/@matematicasdiscretaslibro/cap%C3%ADtulo-3-punto-flotante-c689043db98b>
- [2] [4] [5] <http://puntoflotante.org/formats/fp/>
- [7] <http://puntoflotante.org/formats/fp/>
- [8] <https://www.uv.es/~diaz/mn/node11.html>
- [9] <http://www.portalhuarpe.com.ar/Medhime20/Sitios%20con%20Medhime/Computaci%C3%B3n/COMPUTACION/Menu/modulo%203/paginas/U3-A-SLF-ComaFlotante.htm>