EBook Gratis

APRENDIZAJE floating-point

Free unaffiliated eBook created from
Stack Overflow contributors.

#floating-point

Tabla de contenido

Acerca de

You can share this PDF with anyone you feel could benefit from it, downloaded the latest version from: floating-point

It is an unofficial and free floating-point ebook created for educational purposes. All the content is extracted from Stack Overflow Documentation, which is written by many hardworking individuals at Stack Overflow. It is neither affiliated with Stack Overflow nor official floating-point.

Use the content presented in this book at your own risk; it is not guaranteed to be correct nor accurate, please send your feedback and corrections to info@zzzprojects.com

Capítulo 1: Empezando con el punto flotante

Observaciones

Esta sección proporciona una descripción general de qué es el punto flotante y por qué un desarrollador puede querer usarlo.

También debe mencionar cualquier tema grande dentro de punto flotante y vincular a los temas relacionados. Dado que la Documentación para punto flotante es nueva, es posible que deba crear versiones iniciales de esos temas relacionados.

Examples

Visión general

¿Qué es el punto flotante?

Hay dos tipos de números:

- punto fijo donde un cierto número de dígitos está disponible antes y después del punto de raíz.
- Punto flotante donde hay un cierto número de dígitos disponibles para la mantisa y para el exponente.

Un ejemplo que usa dígitos decimales con tres lugares decimales antes del punto decimal y dos lugares decimales después del lugar decimal:

- 0 se representaría como 000.00
- 0.123 se representaría como 000.12
- 0.00123 estaría representado como 000.00
- 1 se representaría como 001.00
- 1.123 se representaría como 001.12
- 1.00123 se representaría como 001.00
- 123.456 se representaría como 123.45
- 1234.56 es un error porque se almacenaría como 234.56 y eso es simplemente incorrecto

Un ejemplo que usa dígitos decimales con cinco lugares decimales para la mantisa y un lugar decimal para el exponente:

- 0 podría representarse como .00000 x 10 ^ 0
- 0.1 podría representarse como .10000 x 10 ^ 0
- 0.0000123456 podría representarse como .12345 x 10 ^ -4
- 0.000000000123456 podría representarse como .12345 x 10 ^ -9
- 0.00000000001 es un error porque el exponente no es lo suficientemente grande como para

almacenar el número
- 1 podría representarse como .10000 x 10 ^ 1
- 1.123 podría representarse como .11230 x 10 ^ 1
- 1.00123 podría representarse como .10012 x 10 ^ 1
- 123.45678 podría representarse como .12345 x 10 ^ 2
- 123456789.1 podría representarse como .12345 x 10 ^ 9
- 1000000000 es un error porque el exponente no es lo suficientemente grande como para almacenar el número

Por lo tanto, un número de punto flotante puede representar números con magnitudes muy diferentes (0.000000000123456 y 123456789.1) con la misma cantidad de precisión relativa.

Los números de puntos fijos son útiles cuando siempre se necesita un número determinado de lugares decimales, independientemente de la magnitud del número (dinero, por ejemplo). Los números de punto flotante son útiles cuando la magnitud varía y aún se necesita precisión. Por ejemplo: para un ingeniero de caminos, las distancias se miden en metros y .01 de un metro es insignificante, pero para un diseñador de microchips, la diferencia entre 0.0000001 metros y .000000001 metros es enorme, y un físico podría necesitar números enormes y muy, muy Números minúsculos en el mismo cálculo. La precisión en muchas magnitudes diferentes es lo que hace que los números de punto flotante sean útiles.

Cómo funciona

Las computadoras no usan decimal, usan binario y eso causa problemas para el punto flotante porque no todos los números decimales se pueden representar exactamente con un número de punto flotante y eso introduce errores de redondeo en los cálculos.

Después de haber hecho todos los ejemplos en decimal, es importante tener en cuenta que, dado que son binarios, en lugar de almacenar un número de coma flotante como una suma de fracciones decimales:

```
123.875 = 1/10^-2 + 2/10^-1 + 3/10^0 + 8/10^1 + 7/10^2 + 5/10^3
```

las computadoras almacenan números de punto flotante como una suma de fracciones binarias:

```
123.875 = 1/2^-6 + 1/2^-5 + 1/2^-4 + 1/2^-3 + 1/2^-1 + 1/2^0 + 1/2^1 + 1/2^2 + 1/2^3
```

Hay muchas formas diferentes de almacenar patrones de bits que representan esas fracciones, pero la única que la mayoría de las computadoras usan ahora se basa en el estándar IEEE-754. Tiene reglas para almacenar representaciones decimales y binarias y para diferentes tipos de datos de tamaño.

La forma en que se almacenan los números normales utilizando el estándar IEEE es:

- un bit para el signo: almacenado en el MSB, 1 significa negativo y 0 significa positivo
- Algunos bits para el exponente: el sesgo se resta para obtener exponentes positivos y

negativos.
- algunos bits para la mantisa - dígitos después del lugar decimal con un 1 implícito antes del lugar decimal.

Para permitir un subdesbordamiento más gradual, los números desnormalizados (cuando los bits del exponente son todos cero) se tratan de manera especial: el exponente se establece en -126 y el inicio implícito 1 antes de la posición decimal NO se agrega a la mantisa.

Números de punto flotante IEEE-754 de 32 bits

Para un número de punto flotante IEEE-754 normal de 32 bits:

- el bit 32 es el signo
- los bits 24-31 son el exponente - el sesgo es 127
- bits 1-23 son la mantisa

Así que un número normal se calcula como:

```
-1^sign * 2^(exponent-bias) * 1.mantissa
```

Si el patrón de bits fuera:

```
0 10000101 11101111100000000000000
```

Entonces el valor es:

```
-1^0 * 2^(133-127) * 1.111011111
-1^0 *     2^6     * (1 + 1/2 + 1/4 + 1/8 + 1/32 + 1/64 + 1/128 + 1/256 + 1/512)
   1 *      64     * 991/512
         123.875
```

Hay algunos valores especiales:

```
0 11111111 11111111111111111111111 = NaN
0 11111111 00000000000000000000000 = +infinity
1 11111111 00000000000000000000000 = -infinity
0 00000000 00000000000000000000000 = +Zero
1 00000000 00000000000000000000000 = -Zero
```

Los detalles específicos del formato IEEE-754 de 32 bits se pueden encontrar en:

- http://floating-point-gui.de/formats/fp
- https://en.wikipedia.org/wiki/Single-precision_floating-point_format
- https://en.wikipedia.org/wiki/IEEE_floating_point

Lea Empezando con el punto flotante en línea: https://riptutorial.com/es/floating-

point/topic/8816/empezando-con-el-punto-flotante

Creditos

S. No	Capítulos	Contributors
1	Empezando con el punto flotante	Community, Jerry Jeremiah