710193M Arquitectura de computadores II

Aritmética del computador: Representación punto flotante

Facultad de Ingeniería. Universidad del Valle

Marzo de 2022

Contenido

Representación en punto flotante

Definiciones

- En la representación complemento a dos es posible representar enteros positivos o negativos
- Sin embargo, no es posible representar números fraccionarios
- Se requiere un gran número de bits para representar números grandes

Representación en punto flotante

Definiciones

- En el sistema decimal las anteriores limitaciones se superan usando notación científica:
  1. $10000000000 = 1x10^{10}$
  2. $0{,}00000254 = 2{,}54x10^{-6}$
- Esta representación se utiliza para correr la coma decimal, tantas potencias de 10 se le indique. Si se corre la izquierda la potencia es positiva y si se corre a la derecha es potencia negativa.

Representación en complemento a dos

Ejercicio en clase

Transforme las siguientes expresiones en notación científica:

- 15200000000 $\quad 1{,}52 \times 10^{10}$
- 0,0000012445 $\quad 1{,}2445 \times 10^{-6}$
- 0,155457878454 $\quad 1{,}55457878454 \times 10^{-1}$

$\boxed{1{,}554578 \times 10^{-1}}$

Representación en complemento a dos

Ejercicio en clase

Respuesta:

- $1,52x10^{10}$
- $1,2445x10^{-6}$
- 0,155457878454$x10^0$

Limitaciones

Observe que no aplica para todos los casos, como es el caso de 0,155457878454, no se puede acortar la representación utilizando notación científica de lo contrario se podría perder información.

Representación en punto flotante

Definiciones

La técnica de notación científica se puede aplicar a los números binarios.

Notación científica binarios

$$\pm SxB^{\pm E} \tag{1}$$

1. $\pm$ Signo
2. S Mantisa: Parte significativa
3. E Exponente
4. B Base, en binario es 2

Representación en punto flotante

Definiciones

1. Si el bit MSB (más significativo) es 0, el número es positivo y si es 1, el número es negativo
2. El exponente consta de 8 bits.
3. Se utiliza **representación sesgada**

Representación en punto flotante

Representación sesgada

1. Es un valor que se le resta al exponente
2. Es denominado como **sesgo**
3. Tiene un valor de $\boxed{2^{k-1} - 1}$, k es el número de bits del exponente
4. Por lo que en un campo de 8 bits, este comprende entre un valor en el rango -127 a $+128$

0 0 0 0 0 0 0 0 0

1 1 1 1 1 1 1 1 = 255

128 64 32 16 8 4 2 1

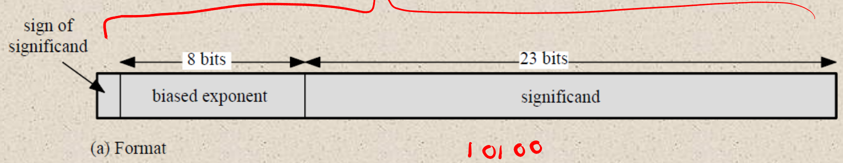
Sesgo $2^7 - 1 = 127$

-127 128

Representación en punto flotante

32

$1\,100 \times 2^{-3}$



10100

$1.1010001 \times 2^{10100} = 0\ 10010011\ 10100010000000000000000 = 1.6328125 \times 2^{20}$

$-1.1010001 \times 2^{10100} = 1\ 10010011\ 10100010000000000000000 = -1.6328125 \times 2^{20}$

$1.1010001 \times 2^{-10100} = 0\ 01101011\ 10100010000000000000000 = 1.6328125 \times 2^{-20}$

$-1.1010001 \times 2^{-10100} = 1\ 01101011\ 10100010000000000000000 = -1.6328125 \times 2^{-20}$

Figura 1: Formato típico de 32 bits en coma flotante

Representación en punto flotante

Observaciones

1. Las siguientes representaciones son equivalentes:
   $0,110 * 2^5$
   $110 * 2^2$
   $0,0110 * 2^6$
   Para simplificarlos cálculos se utiliza la siguiente representación:
   $\pm 0,1bbbb * 2^{\pm E}$
2. Como se puede observar siempre tiene un 1 el el bit más a la izquierda
3. Este bit se puede quitar y cuando se realizan los cálculos se vuelve a colocar

Representación en punto flotante

Observaciones

Por lo que este formato:

1. El signo se almacena en el bit más a la izquierda
2. El primer bit de la parte significativa siempre es 1, por lo que puede quitarse
3. Se suma 127 al exponente original
4. La base es 2

Complemento a dos vs representación en punto flotante

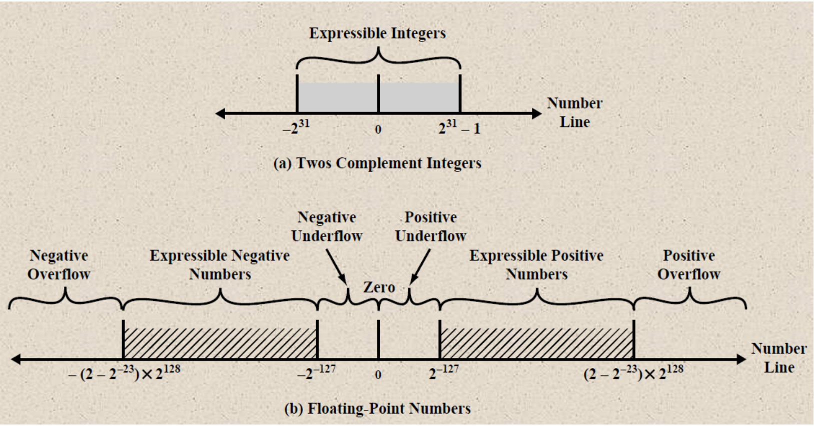


Figura 2: Comparación representación a dos vs punto flotante

Representación en punto flotante

Rango

Para el formato de representación en punto flotante se tienen los siguientes rangos:

1. Números negativos entre $-(2 - 2^{23}) * 2^{128}$ y -2^{-127}
2. Números positivos entre -2^{-127} y $(2 - 2^{23})x2^{128}$

Representación en punto flotante

Regiones

En la recta se pueden observar las siguientes regiones se encuentran excluidas:

1. **Desbordamiento negativo:** números negativos menores que: $-(2-2^{23})x2^{128}$
2. **Agotamiento negativo:** números negativos mayores que: -2^{-127}
3. **Agotamiento positivo:** números positivos menores 2^{-127}
4. **Desbordamiento positivo:** números positivos mayores que $(2-2^{23})x2^{128}$
5. El zero

Representación en punto flotante

Densidad

Se puede observar que el espaciamiento entre los números en coma flotante no es uniforme:

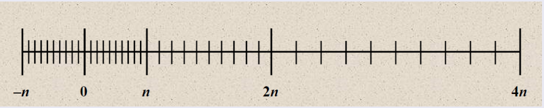


Figura 3: Densidad de los números en coma flotante

Se puede observar que existe una gran población de números cerca al origen

Contenido

Estándar del IEEE para punto flotante

Formatos

1. **Formato simple:** 32 bits
2. **Formato doble:** 64 bits
3. Formato de 128 bits.

Estándar del IEEE para punto flotante

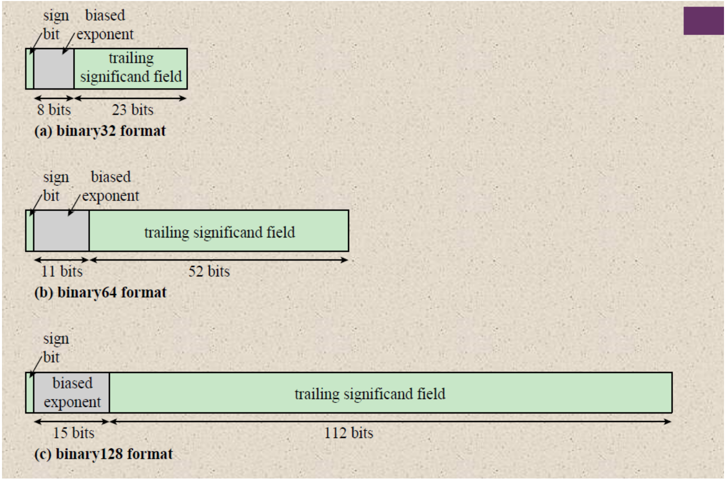


Figura 4: Formatos IEEE 754

Estándar del IEEE para punto flotante

Parameter	Format		
	binary32	binary64	binary128
Storage width (bits)	32	64	128
Exponent width (bits)	8	11	15
Exponent bias	127	1023	16383
Maximum exponent	127	1023	16383
Minimum exponent	–126	–1022	–16382
Approx normal number range (base 10)	$10_{-38}, 10_{+38}$	$10_{-308}, 10_{+308}$	$10_{-4932}, 10_{+4932}$
Trailing significand width (bits)*	23	52	112
Number of exponents	254	2046	32766
Number of fractions	2_{23}	2_{52}	2_{112}
Number of values	$1.98 \leftrightarrow 2_{31}$	$1.99 \leftrightarrow 2_{63}$	$1.99 \leftrightarrow 2_{128}$
Smallest positive normal number	2_{-126}	2_{-1022}	2_{-16362}
Largest positive normal number	$2_{128} - 2_{104}$	$2_{1024} - 2_{971}$	$2_{16384} - 2_{16271}$
Smallest subnormal magnitude	2_{-149}	2_{-1074}	2_{-16494}

Figura 5: Interpretación números en coma flotante

Estándar del IEEE para punto flotante

Ejemplos

Transforme al estándar IEEE 754 de 32 bits los siguientes números:

1. $10{,}5_{10}$
2. 17236_{10}
3. $-127{,}125_{10}$

Estándar del IEEE para punto flotante

Ejemplos

Paso 1: Transforme el número en binario

1. $1010{,}1_2$
2. 1000011010101002_2
3. (Signo negativo) $1111111{,}001_2$

Estándar del IEEE para punto flotante

Ejemplos

Paso 2: Normalice el número

1. $1{,}0101_2 * (10^3)_2$
2. $1{,}00001101010100_2 * (10^{14})_2$
3. (Signo negativo) $1{,}111111001_2 * (10^6)_2$

Estándar del IEEE para punto flotante

Ejemplos

Paso 3: Calcula el exponente, recuerda que para el caso de 32 bits, el sesgo es $2^8 - 1$ entonces es $127_{10} = 1111111_2$

1. $127_{10} + 3_{10} = 130_{10} = 10000010_2$
2. $127_{10} + 14_{10} = 141_{10} = 10001101_2$
3. (Signo negativo) $127_{10} + 6_{10} = 133_{10} = 10000101_2$

Estándar del IEEE para punto flotante

Ejemplos

Paso 4: Se obtiene el número en estándar IEEE 754 de 32 bits:

Signo	Exponente	Mantisa
0	10000010	**0101**0000000000000000000000000000
0	10001101	**00001101010100**000000000000000000
1	10000101	**111111001**00000000000000000000000

Verifiquemos:

http://www.h-schmidt.net/FloatConverter/IEEE754.html

Consultado Feb 2016

Estándar del IEEE para punto flotante

Ejercicio en clase

Transforme al estándar IEEE 754 de 32 bits los siguientes números:

1. $18{,}75_{10} = 10010{,}11_2$
2. $14578{,}5_{10} = 11100011110010{,}1_2$
3. $-0{,}625_{10} = 0{,}101_2$

Estándar del IEEE para punto flotante

Ejercicio en clase

Respuestas:

1. $01000001100101100000000000000000_2$
2. $01000110011000111100101000000000_2$
3. $10111111001000000000000000000000_2$

Estándar del IEEE para punto flotante

Ejercicio en clase

Ahora intenta para 64 bits, recuerda: exponente 11 bits y mantisa 53 bits.

1. $18{,}75_{10} = 10010{,}11_2$
2. $14578{,}5_{10} = 11100011110010{,}1_2$
3. $-0{,}625_{10} = 0{,}101_2$

¿Cuanto es el sesgo?

Recuerda es: $2^{k-1} - 1$, donde k es el número de bits del exponente

Estándar del IEEE para punto flotante

Ejercicio en clase

Respuestas:

1. $0100000000110010110000000000000000000000000000000000000000000000_2$
2. $0100000011001100011110010100000000000000000000000000000000000000_2$
3. $1011111111100100000000000000000000000000000000000000000000000000_2$

Estándar del IEEE para punto flotante

Ejemplos

Transforme a decimal los siguientes números en el estándar IEEE 754 de 32 bits:

1. $01000001001100000000000000000000_2$
2. $11000000110100000000000000000000_2$
3. $01000001001001000000000000000000_2$

Estándar del IEEE para punto flotante

Ejemplos

Paso 1: Reste el sesgo al exponente

1. $10000010_2 - 1111111_2$
   - $130_{10} - 127_{10}$
   - 3_{10}
2. $10000001_2 - 1111111_2$
   - $129_{10} - 127_{10}$
   - 2_{10}
3. $10000100_2 - 1111111_2$
   - $132_{10} - 127_{10}$
   - 5_{10}

Estándar del IEEE para punto flotante

Ejemplos

Paso 2: Agregue 1. a la mantisa

1. $1{,}01100000000000000000000_2$
2. $1{,}10100000000000000000000_2$
3. $1{,}10010000000000000000000_2$

Estándar del IEEE para punto flotante

Ejemplos

Paso 3: Multiplique por 2 elevado al exponente al que se le ha restado el sesgo

1. $1{,}01100000000000000000000_2 * 2^3$
2. $1{,}10100000000000000000000_2 * 2^2$
3. $1{,}10010000000000000000000_2 * 2^5$

En binario, multiplicar por 2 es análogo a multiplicar por 10 en decimal, **sólo debes correr la coma tantas posiciones a la derecha o izquierda**

Estándar del IEEE para punto flotante

Ejemplos

Paso 4: Convierta en decimal, tome en cuenta el bit del signo

1. $1011{,}00000000000000000000_2 = 11_{10}$
2. $-110{,}100000000000000000000_2 = -6{,}5_{10}$
3. $110010{,}000000000000000000_2 = -50_{10}$

Estándar del IEEE para punto flotante

Ejercicio

Transforme a decimal los siguientes números en el estándar IEEE 754 de 32 bits:

1. $01000010000000101000000000000000_2$
2. $11000001011110000000000000000000_2$
3. $01000010110010001000000000000000_2$

Estándar del IEEE para punto flotante

Ejercicio

Respuestas:

1. $33{,}25_{10}$
2. $-15{,}5_{10}$
3. $100{,}25_{10}$

Contenido

Aritmética en punto flotante

Definiciones

- Sumas y restas más complicadas que la multiplicación y división, imagine que tiene que sumar $3 * 10^5 + 4 * 10^3$.
- Por lo tanto, para sumas y restas es necesario ajustar el exponente.

Aritmética en punto flotante

Sumas y restas

Se deben realizar 4 pasos:

1. **Comprobación de cero:** Debido a que la suma y la resta son idénticas, excepto por el cambio de signo, en el caso de restas se cambia el signo del **substraendo**
2. **Ajuste cifras significativas:** Ajuste el menor exponente igual al mayor exponente, realizando los corrimientos de coma necesarios:
   $123 * 10^0 + 456 * 10^{-2} = 123 * 10^0 + 4{,}56 * 10^0 = 127{,}56 * 10^0$
3. **Realice la suma respectiva, tomando en cuenta el signo de los operandos**
4. Normalización, recuerde que la mantisa tiene forma $1.bbbb$

Aritmética en punto flotante

Ejemplo 1

Realice la siguiente operación:

1. $01000010100101010000000000000000_2 +$
   $01000000011100000000000000000000_2$
   $74{,}5_{10} + 3{,}75_{10}$

Aritmética en punto flotante

Ejemplo 1

Paso 1: Comprobación de cero:

1. $01000010100101010000000000000000_2 +$
   $01000000011100000000000000000000_2$

En este caso ya que los dos son positivos y se trata de una suma, no se realiza ningún cambio de signo

Aritmética en punto flotante

Ejemplo 1

Paso 2: Ajuste cifras significativas:

- Exponentes: $10000101_2, 10000000_2$
- Mantisas: $1{,}00101010000000000000000_2, 1{,}11100000000000000000000_2$
- La diferencia entre los exponentes es: $101_2 = 5_{10}$ por lo que se deben correr 5 posiciones a la izquierda la segunda mantisa así:
  Mantisas: $1{,}00101010000000000000000_2$, $0{,}0000111100000000000000000000_2$

Aritmética en punto flotante

Ejemplo 1

Paso 3: Suma: Mantisas normalizadas:
$1{,}00101010000000000000000_2$,
$0{,}0000111000000000000000000000_2$

- Se realiza la suma de las mantisas para obtener: $1{,}00111001000000000000000_2$
- Como en este caso ha quedado de la forma 1,0 no es necesario cambiar el exponente.
- Por lo que el resultado es: $01000010100111001000000000000000_2 = 78{,}25_{10}$.

Aritmética en punto flotante

Ejemplo 2

Realice la siguiente operación:

1. $11000001110011000000000000000000_2 - 01000001100011100000000000000000_2$
   $-25{,}5_{10} - 17{,}75_{10}$

Aritmética en punto flotante

Ejemplo 2

Paso 1: Comprobación de cero:

1. $11000001110011000000000000000000_2 +$
   $11000001100011100000000000000000_2$

En este caso se tiene una resta por lo que el **substraendo** se cambia de signo.

Aritmética en punto flotante

Ejemplo 2

Paso 2: Ajuste cifras significativas:

- Exponentes: $10000011_2, 10000011_2$
- Mantisas: $1{,}10011000000000000000000_2$, $1{,}00011100000000000000000_2$
- Como se puede observar ambos exponentes son iguales, por lo que no se realizan cambios en las mantisas.

Aritmética en punto flotante

Ejemplo 2

textbfPaso 3: Suma: Mantisas: $1{,}10011000000000000000000_2$, $1{,}00011100000000000000000_2$

- Como son ambos negativos, se suma sin problema y se toma en cuenta que el resultado es negativo
- El resultado de la suma es: $10{,}10110100000000000000000_2$
- Como en este caso no ha quedado de la forma 1,0 es necesario correr la coma una posición a la izquierda, **que es equivalente a sumar 1 al exponente** y la mantisa ahora es: $1{,}01011010000000000000000\ \mathbf{0}\ {}_2$
- Por lo que el nuevo exponente es: 10000100_2
- El resultado total es: $11000010001011010000000000000000_2 = -43{,}25_{10}$.

Aritmética en punto flotante

Multiplicación y división

- Para el caso de multiplicación, se suman los exponentes, pero a uno de ellos se le debe restar el sesgo, de lo contrario lo está sumando 2 veces. En el caso de la división se realiza la resta y se le suma el sesgo.
- Las mantisas se multiplican o dividen segundo el caso
- Considere los signos y aplique la **ley de los signos** según el caso

Aritmética en punto flotante

Ejemplo

Realice las siguientes multiplicaciones y divisiones:

1. $01000001100000000000000000000000_2 +$
   $01000000001000000000000000000000_2$
   $16_{10} * 2{,}5_{10}$
2. $01000001100000000000000000000000_2 -$
   $01000000100000000000000000000000_2$
   $16_{10} / 4_{10}$

Aritmética en punto flotante

Ejemplo

Ajuste exponentes:

1. $01000001100000000000000000000000_2 +$
   $01000000001000000000000000000000_2$
   $10000011_2 + 10000000_2 - 1111111_2 = 10000100_2$
2. $01000001100000000000000000000000_2 -$
   $01000000100000000000000000000000_2$
   $10000011_2 - 10000001_2 + 1111111_2 = 10000001_2$

Aritmética en punto flotante

Ejemplo

Ajuste mantisas:

1. $01000001100000000000000000000000_2 +$
   $01000000001000000000000000000000_2$
   $1{,}00000000000000000000000_2 *$
   $1{,}01000000000000000000000 =$
   $1{,}01000000000000000000000_2$
2. $01000001100000000000000000000000_2 -$
   $01000000100000000000000000000000_2$
   $1{,}00000000000000000000000_2 / 1{,}00000000000000000000000 =$
   $1{,}00000000000000000000000_2$

Aritmética en punto flotante

Ejemplo

Uniendo resultados:

1. $01000010001000000000000000000000_2 = 40_{10}$
2. $01000000100000000000000000000000_2 = 4_{10}$

Preguntas

¿Preguntas?

Siguiente clase:
Repertorio de Instrucciones:
Características y funciones