Practica 2

lunes, 28 de septiembre de 2020 02:05

1) Considerando un sistema de Punto Flotante cuya mantisa es fraccionaria, está expresada en BSS con 10 bits y su exponente en BCS con 5 bits, interprete las siguientes cadenas de bits (teniendo en cuenta que la mantisa está codificada en los 10 bits más significativos seguida por los 5 bits que representan al exponente):

010001011101110	111111111111111	111111111100000
000000000100000	000000000000000	100000000000000
000000001110011	000000000011111	000000000111111

$0,M * B^E$ $B=2$

$X, \overbrace{0\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 1}^{M}\quad \overbrace{0\ 1\ 1\ 1\ 0}^{E}$ BCS=NORMALIZADA

$2^{-2}\quad 2^{-6}\quad 2^{-8}\quad 2^{-9}\quad 2^{-10}\qquad 14$

$V = M * B^E$

$(2^{-2}+2^{-6}+2^{-8}+2^{-9}+2^{-10}) * 2^{14}$

$2^{12}+2^{8}+2^{6}+2^{5}+2^{5}$

0 0 0 0 0 0 0 0 0 1 , 0 0 0 0 0

$2^{-10} \times 2^{0}$

2^{-10}

0 0 0 0 0 0 0 0 1 1 , 1 0 0 1 1

$2^{-9}+2^{-10}$ -3

$(2^{-9}+2^{-10}) * 2^{-3}$

$2^{-12}+2^{-13}$

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

-15

$(2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5}+2^{-6}+2^{-7}+2^{-8}+2^{-9}+2^{-10}) \times 2^{-15}$

$2^{-16}+2^{-17}+2^{-18}+2^{-19}+2^{-20}+2^{-21}+2^{-22}+2^{-23}+2^{-24}+2^{-25}$

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

$0 \times 2^{0} = 0$

0 0 0 0 0 0 0 0 0 0 1 1 1 1 1

0 -15

0×2^{-15}

1 1 1 1 1 1 1 1 1 1 0 0 0 0 0

$2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5}+2^{-6}+2^{-7}+2^{-8}+2^{-9}+2^{-10}$

1 0 0 0 0 0 0 0 0 0 0 0 0 0 0

2^{-1} 0

$2^{-1} * 2^{0} = \frac{1}{2}$

0 0 0 0 0 0 0 0 0 1 1 1 1 1 1

2^{-10} 2^{-15}

$*$

2^{-25}

$\overbrace{0\ 1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 1}^{M}, \overbrace{0\ 1\ 1\ 1\ 0}^{E}$

279 14 279×2^{14}

0 0 0 0 0 0 0 0 0 1 , 0 0 0 0 0

1 0

1×2^{0}

0 0 0 0 0 0 0 0 1 1 , 1 0 0 1 1

3 -3

3×2^{-3}

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

1023 -15

1023×2^{-15}

0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

0 0×2^{0} 0

0 0 0 0 0 0 0 0 0 0 1 1 1 1 1

0 (0×2^{-15}) -15

1 1 1 1 1 1 1 1 1 1 0 0 0 0 0

1023 (1023×2^{0}) 0

1 0 0 0 0 0 0 0 0 0 0 0 0 0 0

512 (512×2^{0}) 0

0 0 0 0 0 0 0 0 0 1 1 1 1 1 1

1 (1×2^{-15}) -15

2) Interprete las mismas cadenas del ejercicio 1 pero ahora considerando que están representadas en un sistema de Punto Flotante cuya mantisa es fraccionaria, normalizada, codificada en BCS de 10 bits y su exponente, en BSS de 5 bits. Identifique aquellas cadenas que no pueden ser interpretadas o presentan algún tipo de problema. Describa el motivo.

Bcs M=Mantisa F=Fraccionaria N=Normalizada

$X, 0 \overbrace{1\ 0\ 0\ 0\ 1\ 0\ 1\ 1\ 1}^{MFN}\quad \overbrace{0\ 1\ 1\ 1\ 0}^{E}$

$+\ 2^{-1}\quad 2^{-5}\quad 2^{-7}\ 2^{-8}\quad 2^{-9}\qquad 14$

$(2^{-1}+2^{-5}+2^{-7}+2^{-8}+2^{-9}) \cdot 2^{14}$

$2^{13}+2^{9}+2^{7}+2^{6}+2^{5}$

0 0 0 0 0 0 0 0 0 1 , 0 0 0 0 0

$2^{-9} \times 2^{0} =$ No se puede poner porque se sale del rango y es Irrepresentable en normalización ya que el exponente queda en negativo

0 0 0 0 0 0 0 0 1 1 , 1 0 0 1 1

$2^{-8}\ 2^{-9}$ 19

$(2^{-8}+2^{-9}) * 2^{19} = 2^{11}+2^{10} = 2048+1024$

0 1 1 0 0 0 0 0 0 0 X X X X X

$(2^{-1}+2^{-2}) * 2^{12} = 2^{11}+2^{10}$ Busco un exponente para que se mantenga la igualdad

0 1 1 0 0 0 0 0 0 0 1 1 1 0 0

$2^{-1}\ 2^{-2}$ 12

0 0 0 0 0 0 0 0 0 1 1 1 1 1 1

2^{-9} 31

$2^{-9} \cdot 2^{31} = 2^{22}$

$2^{-1} \cdot 2^{23} = 2^{22}$

0 1 0 0 0 0 0 0 0 0 1 0 1 1 0

2^{-1} 22

3) Repita el ejercicio 2 considerando una la mantisa fraccionaria normalizada con bit implícito.

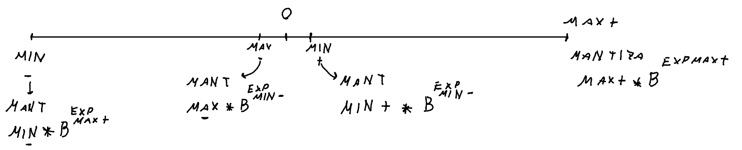
Ya que hay bit implícito indica que empieza con 0,1 entonces ya me deshago de el problema de

normalizar, ya que todos están normalizados.
Antes la mantisa fraccionaria empezaba como 0,mantiza. Ahora es 0,1 mantisa

4) Calcule el rango y la resolución en los extremos inferior negativo, superior negativo, inferior positivo y superior positivo para los siguientes sistemas de representación en punto flotante:

a) Mantisa fraccionaria en BSS de 8 bits y exponente en BSS de 4 bits.
b) Mantisa fraccionaria normalizada en BSS de 15 bits y exponente en CA1 de 10 bits.
c) Mantisa fraccionaria normalizada en BSS de 15 bits y exponente en BSS de 10 bits.
d) Mantisa fraccionaria normalizada con bit implícito en BSS de 15 bits y exponente en CA2 de 7 bits.
e) Mantisa fraccionaria normalizada con bit implícito en BCS de 15 bits y exponente en Exceso de 5 bits.

- En las mantisas BSS no se puede expresar números negativos, con lo que aun con exponente negativo expresaremos un número positivo por un factor de escala menor a 1, pero también positivo. Ejemplo: $2 \times 2^{-4} = 0{,}125$.
- Las mantisas fraccionarias suponen el punto al principio de la mantisa.
- Los exponentes negativos indican factores de escala menores a 1 que mejoran la resolución.
- Mantisa normalizada implica que empieza con 1, o sea mantisa mínima 0,1 para la fraccionaria, igual a 0,5 en decimal. Esto hace que no se pueda representar el 0.
- Mantisa normalizada con bit implícito, significa agregar un 1 al principio de la misma al interpretarla. Ejemplo: 00000 se interpreta 0,100000, o 0,5 en base 10.



A) M.F.BSS 0, $\underbrace{1\ 1\ 1\ 1\ 1\ 1\ 1\ 1}$ $\underbrace{1\ 1\ 1\ 1}_{15}$ BSS

$(2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5}+2^{-6}+2^{-7}+2^{-8}) * 2^{15}$ No hay negativos

M = 0,11111111
M = -1,00000000
0,00000001

(MAX+) $2^{14}+2^{13}+2^{12}+2^{11}+2^{10}+2^{9}+2^{8}+2^{7} \Leftrightarrow (1-2^{-8}) * 2^{15} = 2^{15}-2^{7}$

MIN ⟶ 0,00 0 0 0 0 0 1 $\underbrace{0\,0\,0\,0}$
$2^{-8} \cdot 2^{0} = 2^{-8}$

RANGO
$(2^{-8}), (2^{15}-2^{7})$

RESOLUCIÓN
(2^{-8})

R2
$(2^{-8} \cdot 2^{15}) = 2^{7}$

b. Mantisa fraccionaria normalizada en BSS de 15 bits y exponente en CA1 10 bits

Como la mantisa es BSS, los números representables son todos positivos. Para obtener el mínimo, minimizo mantisa (normalizada) y exponente, el cual debe ser el máximo negativo en ca1.

$N_1 = 0{,}100000000000000 \times 2^{1000000000} = 2^{-1} \times 2^{-511} = 0{,}5 \times 2^{-511}$

Se observa que el cero no es representable en este sistema, por estar la mantisa normalizada.

Para obtener el máximo, maximizo la mantisa (en BSS) y el exponente (en Ca1),

$N_2 = 0{,}111111111111111 \times 2^{0111111111} = (1- 2^{-15}) \times 2^{+511}$

Entonces el rango es: $= [N_1, N_2] =$ **$[0{,}5 \times 2^{-511}; (1- 2^{-15}) \times 2^{+511}]$**

Para la resolución en el extremo inferior, debo obtener el número siguiente al mínimo, lo llamos N3; para ello tomo la siguiente mantisa y utilizo el mismo exponente que para representar ese número

$N_3 = 0{,}100000000000001 \times 2^{1000000000} = (2^{-1} + 2^{-15}) \times 2^{-511}$

R_1= Resolución extremo inferior $= N_3 - N_1 = (0{,}100000000000001- 0{,}100000000000000) \times 2^{1000000000}$

$= 0{,}000000000000001 \times 2^{1000000000} = \mathbf{2^{-15} . 2^{-511}}$

Para la resolución extremo superior debemos obtener el número representable en este sistema anterior al más grande, lo llamaremos N4. Para ello, debemos maximizar mantisa al número anterior a la mantisa máxima y maximizar el exponente:

$N_4 = 0{,}111111111111110 \times 2^{0111111111}$

R_2= Resolución extremo superior $= N_2 - N_4 = (0{,}111111111111111 - 0{,}111111111111110) \times 2^{0111111111}$

$= 0{,}000000000000001 \times 2^{0111111111} = \mathbf{2^{-15} \times 2^{+511}}$

c. Mantisa fraccionaria normalizada con bit implícito en BCS de 15 bits y exponente en Exceso 5 bits

En este sistema tenemos valores positivos y negativos.
Para obtener el mínimo positivo (N1), colocamos la mantisa mínima positiva (primer bit en cero por ser BCS), pero asumiendo un digito en 1 de más después de la coma por estar normalizada con bit implícito:
Exponente en exceso = 10000 + 10000 = 00000 (-16 en exceso)

$N_1 = 0\ 0{,}100000000000000 \times 2^{00000} = + (2^{-1}) \times 2^{-16} =$ **$+ 0{,}5 \times 2^{-16}$** (el más chico +)

Para obtener el máximo positivo (N2), maximizo mantisa (recordando bit implícito) y exponente

$N_2 = 0\ 0{,}111111111111111 \times 2^{11111} =$ **$+ (1 - 2^{-15}) \times 2^{+15}$** (el más grande +)

Para la resolución en el extremo inferior obtengo el número positivo más cercano a N1

$N_3 = 0\ 0{,}100000000000001 \times 2^{00000} = + (2^{-1} + 2^{-15}) \times 2^{-16}$ (el más cercano a N_1)

$R_1 = N_3 - N_1 = (2^{-1} + 2^{-15} - 2^{-1}) \times 2^{-16} = \mathbf{(2^{-15}) \times 2^{-16} = 2^{-31}}$

Para la resolución en el extremo superior obtengo el número positivo más cercano a N2

$N_4 = 0\ 0{,}111111111111110 \times 2^{11111}$ (el más cercano a N_2)

$R_2 = N_2 - N_4 = (0{,}111111111111111-0{,}111111111111110) \times 2^{+15} = 2^{-15} \times 2^{+15} = 2^{0} = 1$

Como tenemos números negativos debemos calcular los valores negativos mínimo y máximo:

$N_5 = 1\ 0{,}100000000000000 \times 2^{00000} = - (2^{-1}) \times 2^{-16} = - 0{,}5 \times 2^{-16}$ (mínimo negativo en módulo)

$N_6 = 1\ 0{,}111111111111111 \times 2^{11111} = - (1 - 2^{-15}) \times 2^{+15}$ (máximo negativo en módulo)

En este caso, los valores que se obtienen para las resoluciones en los extremos mínimo y máximo negativo son los mismos que los positivos, en módulo.

El Rango es $[N_6, N_5] \cup [N_1, N_2]$ U: Unión

Rango : $[- (1 - 2^{-15}) \times 2^{+15}; - 0{,}5 \times 2^{-16}] \cup [+ 0{,}5 \times 2^{-16}; + (1 - 2^{-15}) \times 2^{+15}]$

c) MAX(+) 0,1111111111111111 1111111111

MAX(+) = $(2^{-15}-1)$: MANTIZA MAX 2047

N°MAX(+) = $(2^{-15}-1) \times 2^{2047}$

MIN 0,000000000000001 0000000001

NMIN(+) $2^{-15} \cdot 2^{1} = 2^{-14}$

RANGO
$(2^{-14}), ((2^{-15}-1) \times 2^{2047})$

RESOLUCIÓN
(2^{-15})

d) MANT. FRA. NORM. bit IMPLICITO, BSS. 15 bits
EXP. Car. 7 bits

N°Max 0,1 111111111111111 0111111
$(2^{-16}-1) \cdot 2^{63} = 2^{47} - 2^{63}$ 63

N°MIN 0,1 000000000000000 1011 → -8
$2^{-1} \times 2^{-8} = 2^{-9}$

RANGO
$(2^{-9}), (2^{47}-2^{63})$

RESOLUCIÓN
(2^{-9})

e) M.F.N. Bi. Bcs. 15b. Al ser BCS y bit implícito, siempre va a ser negativo y por lo tanto no tengo positivos
Exp. Ex 5 bits

N°Max(-) 0,1 000000000000000 10001
$-(2^{-1}) \times 2^{1} = -1$ Ex = 00001

N°MIN(-) 0,1 111111111111111 00000
$(2^{-16}-1) \times 2^{-16} = (2^{-32}-2^{-16})$ −16

RANGO
$(-(2^{-32}-2^{-16})), -1$

RESOLUCIÓN
$(-(2^{-32}-2^{-16}))$

5) Para cada sistema del ejercicio anterior obtenga todas las representaciones posibles para los siguientes números:

a) El número mínimo y el número máximo.
b) El máximo negativo y el mínimo positivo.
c) `0 ; 1 ; 9 ; -5,0625 ; 34000,5 ; 0,015625`

0 : a) 0,00000000 0000 No se pueden representar números negativos en este sistema

B) No se puede

C) No existe el 0 de forma normalizada no esta en el rango

D) No existe el 0 de forma normalizada y el numero mas chico que se podria reprecentar es el 0,5

E) Idem a "D"

1 : a) 0,10000000 0010 → $2^{-1} \times 2^{2} = 2^{1} = 1$

b) 0,1000000000000000 0000000010 → $2^{-1} \times 2^{1} = 2^{0} = 1$

c) 0,100000000000000 000000001 → $2^{-1} \times 2^{1} = 1$

d) 0,11000000000000 1000000011 → $-2^{-3} \times 2^{3} = -2^{0} = 1$

No se puede representar el 1 en el d, ya que al tener bit implícito, siempre van a ser números negativos

9 : A) 0,10000000 1001 → $2^{-1} \times 2$

B) 0,1001000000000000 0000000100 → $(2^{-1}+2^{-4}) \cdot 2^{4} = 2^{3} + 2^{0} \to 8+1=9$

c) 0,100100000000000 000000100 → $(2^{-1}+2^{-4}) \cdot 2^{4} = 1+8 = 9$

D) 0,1 00100000000000 0000000100 → $(2^{-1}+2^{-4}) \cdot 2^{4} = 1+8 = 9$

Regla, para pasar cualquier numero, busco el numero para que se multiplique por 2 y si no me alcanza
busco una forma para que se cancelen y pueda generar ese numero con una suma

e) 0,1 00100000000000 10000

No se puede representar el 9 ya que al ser bit implícito siempre es negativo

−5,0625: a) NO SE PUEDE → BSS

B) IDEM → A

C) IDEM → A

D IDEM → A

E) 0,1 010001000000000 10011 → $-(2^{-1}+2^{-3}+2^{-7}) \times 2^{3} =$
$2^{2}+1+2^{-4} = 4+1+0,0625$
$= -5,0625$

34000,5 : Se encuentra fuera de rango que todas las opciones

0,015625: a) 0,00000100 0000 = $2^{-6} = 0,015625$
2^{-6}

B) 0,10000000 1010 = $2^{-1} \times 2^{-5} = 2^{-6}$

C) 0,10000100 0001 = $2^{-1} + 2$ ≠ R

d) 0,1

e)

6) Describa como influyen las siguientes variantes en el rango y la resolución de un sistema:

a) Mantisa con signo y sin signo.
b) Exponente con signo y sin signo.
c) El tamaño de la mantisa.
d) El tamaño del exponente.
e) Que la mantisa sea fraccionaria, fraccionaria normalizada y fraccionaria normalizada con bit implícito.

7) Efectúe las siguientes sumas en un sistema de punto flotante con mantisa en BSS de 8 bits y exponente en BCS 8 bits.

```
  00001111 00000011       01111111 00000000       00000001 00000111
+ 00001000 00000010     + 11111100 10000001     + 00011100 00000000
  ???????? ????????       ???????? ????????       ???????? ????????
```

+ 00001111 00000011 + 15×2^{3}
00001000 00000010 $8 \times 2^{2} \to 4 \times 2^{3}$

SE MANTIENE

+ 00001111 00000011 → 15×2^{3} +
00000100 00000011 → 4×2^{3}
00010011 → 19×2^{3} →

B)

```
 0 1 1 1 1 1 1 1   0 0 0 0 0 0 0 0     127 × 2^0
 1 1 1 1 1 1 1 0   1 0 0 0 0 0 0 1     254 × 2^-1 → 126 × 2^0
                                        254/2 = 126 × 1
 0 1 1 1 1 1 1 1   0 0 0 0 0 0 0 0  → 127 × 2^0
 0 1 1 1 1 1 1 0   0 0 0 0 0 0 0 0  → 126 × 2^0
 ---------------
 1 1 1 1 1 1 0 1  → 253 × 2^0 ✓
```

c)

```
 0 0 0 0 0 0 0 1   0 0 0 0 0 1 1 1     1 × 2^7 → 128 × 2^0
 0 0 0 1 1 1 0 0   0 0 0 0 0 0 0 0     28 × 2^0

 1 0 0 0 0 0 0 0   0 0 0 0 0 0 0 0
 0 0 0 1 1 1 0 0   0 0 0 0 0 0 0 0
 -----------------------------------
 1 0 0 1 1 1 0 0   156 ~~× 2^0~~ → 156
```

8) Obtenga las representaciones para los números 8,625; 0,4 y 2,5 en los siguientes sistemas. En caso que no sea posible hacerlo exactamente, represente la mejor aproximación posible.

   a) Mantisa fraccionaria normalizada en BSS de 5 bits y exponente en Ca2 de 4 bits.
   b) Mantisa fraccionaria normalizada en BCS de 10 bits y exponente en Ca2 de 3 bits.

$8,625 = 1000,101 \times 2^0 \rightarrow 0,1000101 \times 2^4$ Al solo tener 5 bits no se puede representar Pero vamos a buscar un valor aproximado

$0,1000101 \xrightarrow{5 bits} 0,10001 \times 2^4$

$(2^{-1} + 2^{-5}) \times 2^4 \rightarrow 2^3 + 2^{-1} = 8,5$

ERROR1 = 8,625 − 8,5 = 0,125 ✓

$0,10010 \times 2^4 \rightarrow (2^{-1}+2^{-4}) \times 2^4 = 2^3+1 = 8+1 = 9$

ERROR2 = 9 − 8,5 = 0,5

Me quedo con el error mas bajo

B) $01000,101 \times 2^0 = 0\ 0,100010100 \times 2^4$

El exponente 4 no se puede expresar con 3 bits en ca2

$00,111111111 + 2^3 = (1-2^9) \times 8$

EL MAS CERCANO

0011 111111111

1) 0,4 M.F.N. 5B.BSS EXP.4b.CA2

$0,4 \times 2 = 0,8$
$0,8 \times 2 = 1,6$ $\quad 0,4 = 0,01100 \times 2^0$
$0,6 \times 2 = 1,2$ $\quad 0,11001 \times 2^{-1} = 0,11001 \times 2^{1111} = 0,390625$
$0,2 \times 2 = 0,4$ $\quad -1\ CA_2 = 1111$
$0,4 \times 2 = 0,8$ $\quad$ EL QUE SIGUE
$0,8 \times 2 = 1,6$ $\quad 0,11010 \times 2^{-1} \times 2^{-1} = 0,40625$

ERROR1: 0,4 − 0,390625 = 0,009375

ERROR2: 0,40625 − 0,4 = 0,00625 → EL MENOR

B) M.F.N. 10B. BCS EXP 3bits CA2

$0,4 \times 2 = 0,8$
$0,8 \times 2 = 1,6$ $\quad 0,4 = 00,011001100 \times 2^0$
$0,6 \times 2 = 1,2$ $\quad$ NOR: $00,110011001 \times 2^{-1}$
$0,2 \times 2 = 0,4$ $\quad (2^{-1}+2^{-2}+2^{-5}+2^{-6}+2^{-9}) \times 2^{-1}$
$0,4 \times 2 = 0,8$ $\quad 2^{-2}+2^{-3}+2^{-6}+2^{-7}+2^{-10} = 0,399414062$
$0,8 \times 2 = 1,6$ $\quad$ EL QUE LE SIGUE
$0,6 \times 2 = 1,2$ $\quad 00,110011010 \times 2^{-1}$
$0,2 \times 2 = 0,4$ $\quad (2^{-1}+2^{-2}+2^{-5}+2^{-6}+2^{-8}) \times 2^{-1} = 0,400390625$
$0,4 \times 2 = 0,8$
$0,8 \times 2 = 1,6$

ERROR1 = 0,4 − 0,399414062 = 0,000585938

ERROR2: 0,400390625 − 0,4 = 0,000390625

→ MENOR ERROR

2,5

a) M.F.N. 5B.BSS EXP. 4B.CA2

$2,5 = 10,1 \times 2^0 = 0,101 \times 2^2 = 0,10100 \times 2^{0010}$

$(2^{-1}+2^{-3}) \times 2^2 = 2^1 + 2^{-1} = 2 + 0,5 = 2,5$

0010 10100

b) M.F.N. 10b. BCS EXP. 3b. CA2

$2,5 = 00,101000000 \times 2^2 = (2^{-1}+2^{-3}) \times 2^2 = 2 + 0,5 = 2,5$

0,1010000000 010

9) Definimos de la siguiente forma el Error Absoluto ($EA(x)$) y Error Relativo ($ER(x)$) de un número x en un sistema de numeración:

$$EA(x) = |x_R - x| \qquad ER(x) = \frac{EA(x)}{x}$$

8,625

a) M.F.N. 5b. BSS

$8,5 < 8,625 < ((2^{-1} + 2^{-5}) \times 2^4 = 9)$

EA = 8,625 − 8,5 = 0,125 (menor error)

$E_R = E_A / N° \text{ REPRESENTAR} \longrightarrow 0,125/8,625 \sim 0,0145$

b) En este caso se obtuvo que el numero mas cercano al representable es el 7,984775

$E_A = 8,625 - 7,984775 = 0,640625$

$E_R = E_A / N° \longrightarrow 0,640625/8,625 \sim 0,0742753623$

0,4

a) $0,390625 < 0,4 < 0,40625$

$E_A = 0,40625 - 0,4 = 0,00625$ (MENOR ERROR)

$E_R = 0,000390625/0,4$

2,5

a) $E_A = 0$

b) $E_A = 0$

10) Considerando que en los procesos de truncamiento o redondeo la elección se basa en la representación mas cercana, estime el Error Absoluto Máximo (EA_{MAX}) cometido en las representaciones del ejercicio 1. Recuerde que la distancia entre 2 representaciones sucesivas se conoce como resolución (R), por lo que:

$$EA_{MAX}(x) \leqslant \frac{R}{2}$$

11) Tome un sistema de punto flotante cualquiera y dibuje en un gráfico como varía el Error Absoluto y el Error Relativo en función del número que se quiere representar.

El IEE 754 es un estándar de aritmética en coma flotante. Este estándar significan como deben representarse los números en coma flotante con simple precisión (32 bits) o doble precisión (64 bits), y también cómo deben realizarse las operaciones aritméticas con ellos .

Emplea mantisa fraccionaria, normalizada y en representación signo magnitud (M y S), sin almacenar el primer digito, que es igual a 1. El exponente se representa en exceso, que en este caso no se toma como 2n-1, sino como 2n-1-1

Simple Precisión:
El estándar IEEE-754 para la representación en simple precisión de números en coma flotante exige una cadena de 32 bits. El primer bit es el bit de signo (S), los siguientes 8 son los bits del exponente "E" y los restantes 23 son la mantisa de (M)

Doble Precisión:
El estándar IEEE-754 para la representación de números en coma flotante que exige una cadena de 64bits. El primer bit es el bit de signo (S), los siguientes 11 son los bits del exponente "E" y los restantes 52 son mantisa (M)

13) ¿Qué valores representan las siguientes cadenas binarias codificadas en el estándar IEEE 754?

a) `0 11000100 00000000000000000000000`
b) `1 11111110 10100000000000000000000`
c) `0 00000000 00000000000000000000001`
d) `0 00000000 10011000000000000000000`
e) `1 00000000 00000000000000000000000`
f) `0 11111111 00000000000000000000000`
g) `0 11111111 00000100000000000000000`

h) `0 01100010100 0000000000000000000000000000000000000000000000000000`
i) `0 11010101110 1010000000000000100000000000000000000000000000000000`
j) `0 00000000000 0101000000000000000000000000000000000000000000000000`
k) `1 11111111111 1111100000000000000000000000000000000000000000000000`

$\overbrace{0}^{\text{SIGNO}}$ $\overbrace{11000100}^{\text{EXPONENTE}}$ $\overbrace{00000000000000000000000}^{\text{MANTIZA}}$

1 bit   8 bits   23 bits

EXP: 8 bits $\longrightarrow$ Ex $128 - 1 = 127$

RESTAMOS EL Ex

$11000100 \longrightarrow (128+64+4) - 127 = 196 - 127 = 69$

MANTISA: Se antepone un 1, y el valor en este caso es 1,0

La cadena de representación es: $+1,0\,2^{+69}$

$\overbrace{1}^{-}$ $\overbrace{11111110}^{E}$ $\overbrace{10100000000000000000000}^{M}$

$E = 254$ $2^{-1}\ 2^{-3} \longrightarrow 0,5 + 0,125 = 0,625$

$= 254 - 127 = 127$

$M = M + 1 \longrightarrow 0,625 + 1 = 1,625$

$-M \times 2^{E}$

$-1,625 \times 2^{127}$

c) 0   00000000   00000000000000000000001

DESNORMALIZADA

$(-1)^0 \times 2^{23} \times 2^{-126}$

$\cancel{1\times}\ 2^{23} \times 2^{-126} = 2^{-103}$

D) 0   00000000   10011000000000000000000

$(-1)^0 \times (2^{-1} + 2^{-4} + 2^{-5}) \times 2^{-126}$

$2^{-127} + 2^{-130} + 2^{-131}$

e) 1 00000000 00000000000000000000000

-0

F) 0 11111111 00000000000000000000000

$\cancel{(-1)}^{signo} \times \cancel{1},0 \times 2^{256-127} \rightarrow 2^{129}$

G)

0 11111111 00000100000000000000000

$1+2^{-6} \times 2^{256-127} \rightarrow (1+2^{-6}) \times 2^{129} = 2^{129} + 2^{123}$

H) 0 01100010100 0000000000000000000000000000000000000000000000000000

512+256+16+4

788

$\cancel{(-1)^{0} \times 0 \times 2^{-1022}} = 0$

I) 0 11010101110 1010000000000000100000000000000000000000000000000000

1024+512+160+14 $\quad 2^{-1} + 2^{-3} + 2^{-17}$

1536+14+160

1550+160 = 1710

$(-1)^{0} \times (1 + 2^{-1} + 2^{-3} + 2^{-17}) \times 2^{1710-127}$

$(1 + 2^{-1} + 2^{-3} + 2^{-17}) \times 2^{1583}$

J) 0 00000000000 0101000000000000000000000000000000000000000000000000

$(-1)^{0} \times (2^{-2} + 2^{-4}) \times 2^{0-2047}$

$(2^{-2}+2^{-4}) \times 2^{-2047}$

K) 1 11111111111 1111100000000000000000000000000000000000000000000000

$(-1)^{1} \times (2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5}) \times 2^{\cancel{2047-2047}}$

$-(2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5})$

14) Hallar la representación de los siguientes números en simple y en doble precisión:

1 ; 13 ; 257 ; -40000 ; 0,0625

1: 01,00000....000 × 2^{0} = 01,00000...000 × $2^{01111111}$

EXP = 0 + 127 = 01111111 (0 en exceso 127)

$\overset{signo}{0}$ $\overbrace{01111111}^{EXP}$ $\overbrace{00000000000000000000000}^{M}$

13: 01101,00000 000 × 2^{0}

01,$\underbrace{101000}_{M}$,..000 × 2^{3}

0,1,101000 000 × $2^{10000010}$

EXP = 3+127 = 130 = 10000010 (3 en ex 127)

$\underset{S}{0}$ $\underset{E}{10000010}$ $\underset{M}{10100000000000000000000}$

$(-1)^{0} \times (1 + 2^{-1} + 2^{-3}) \times 2^{130-127}$

$(1 + 2^{-1} + 2^{-3}) \times 2^{3} \rightarrow 2^{3} + 2^{2} + 1 \rightarrow 8+4+1 = 13$

257 $\quad$ 00100000001,000...00 × 2^{0}

001,00000001 000.00 × 2^{8}

EXP = 8 + 1023 = 1031

0 10000000111 0000000100000000000000000000000000000000000000000000

1024 + 7 = 1031

$(-1)^{0} \times (1 + 2^{-8}) \times 2^{1031-1023}$

$(1+2^{-8}) \times 2^{8} = 2^{8} + 2^{0} = 256+1 = 257$

⊖40000 = 1 1001110001000000 × 2^{0} →

S M → 1 1,00111000100000000000000 × 2^{15}

E $\quad$ EXP = 15 + 127 = 142 → 10001110

S E M

-40000 = 1 10001110 00111000100000000000000

0,0625 = $\overset{S}{0}$ 0,00010000.....0 × 2^{0}

= 1,00000 2^{4}

EXP: 4 + 127 = 131 = 10000011 = (4 en exceso 127)

entonces

0 10000011 00000000000000000000000

15) Calcule el rango y la resolución en los extremos inferior negativo y superior positivo para los sistemas de simple precisión y doble precisión tales como se los define en el estándar IEEE 754. ¿Cuál es el menor número positivo distinto de 0 que se puede representar?

RANGO

$(-1)^{0} \times 1,M \times 2^{EXP-127}$ $\quad$ EXP = 254

M = TODOS "1"

$\underset{S}{0}$ $\underbrace{11111110}_{EXP}$ $\underbrace{11111111111111111111111}_{M}$

$(-1)^{0} \times (2^{8}-2) \times 2^{254-127}$

$(2^{8}-2) \times 2^{127}$

$2^{135} - 2^{127}$

$\underbrace{1}_{s}\ \underbrace{11111110}_{Exp}\ \underbrace{11111111111111111111111}_{M}$

16) ¿Por que motivo la mantisa no está normalizada se cuando el exponente es `0` y la mantisa no es nula?