Organización de Computadoras
TP 3: Punto Flotante

Curso 2021

Prof. Jorge Runco

$$a)\ M\begin{cases} fraccionaria \\ BSS\ 6\ bits \end{cases} \qquad E\begin{cases} BCS \\ 4\ bits \end{cases}$$

$$\underbrace{010111}_{M}\ \underbrace{0110}_{E} \Rightarrow 0{,}010111\ x\ 2^{0110} = (2^{-2}+2^{-4}+2^{-5}+2^{-6})\ x\ 2^{6} = 2^{4}+2^{2}+2^{1}+2^{0} =$$

$$= 16+4+2+1 = 23 \quad \Leftarrow$$

$$\underbrace{000001}_{M}\ \underbrace{0000}_{E} \Rightarrow 0{,}000001\ x\ 2^{0000} = (2^{-6})\ x\ 2^{0} = 2^{-6} \quad \Leftarrow$$

$$\underbrace{000011}_{M}\ \underbrace{1001}_{E} \Rightarrow 0{,}000011\ x\ 2^{1001} = (2^{-5}+2^{-6})\ x\ 2^{-1} = 2^{-6}+2^{-7} \quad \Leftarrow$$

$$\underbrace{111111}_{M}\ \underbrace{1111}_{E} \Rightarrow 0{,}111111\ x\ 2^{1111} = (2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5}+2^{-6})\ x\ 2^{-7} =$$

$$= 2^{-8}+2^{-9}+2^{-10}+2^{-11}+2^{-12}+2^{-13} \quad \Leftarrow$$

$$\underbrace{000000}_{M}\ \underbrace{0000}_{E} \Rightarrow 0{,}000000\ x\ 2^{0000} = 0\ x\ 2^{0} = 0 \quad \Leftarrow$$

$$\underbrace{000000}_{M}\ \underbrace{1111}_{E} \Rightarrow 0{,}000000\ x\ 2^{1111} = 0\ x\ 2^{-7} = 0 \quad \Leftarrow$$

$$\underbrace{111111}_{M}\ \underbrace{0000}_{E} \Rightarrow 0{,}111111\ x\ 2^{0000} = (2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5}+2^{-6})\ x\ 2^{0} =$$
$$= 2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5}+2^{-6} \quad \Leftarrow$$

$$\underbrace{100000}_{M}\ \underbrace{0000}_{E} \Rightarrow 0{,}100000\ x\ 2^{0000} = 2^{-1}\ x\ 2^{0} = 0{,}5 \quad \Leftarrow$$

$$\underbrace{000001}_{M}\ \underbrace{1111}_{E} \Rightarrow 0{,}000001\ x\ 2^{1111} = 2^{-6}\ x\ 2^{-7} = 2^{-13} \quad \Leftarrow$$

$$a)\ M\begin{cases} fraccionaria \\ BCS\ 6\ bits \end{cases} \qquad E\begin{cases} BCS \\ 4\ bits \end{cases}$$

$$\underbrace{0}_{S}\ \underbrace{10111}_{M}\ \underbrace{0110}_{E} \Rightarrow 0\ 0{,}10111\ x\ 2^{0110} = +(2^{-1}+2^{-3}+2^{-4}+2^{-5})\ x\ 2^{6} = +(2^{5}+2^{3}+2^{2}+2^{1}) =$$
$$= 32+8+4+2 = 46 \quad \Leftarrow$$

$$\underbrace{0}_{S}\ \underbrace{00001}_{M}\ \underbrace{0000}_{E} \Rightarrow 0\ 0{,}00001\ x\ 2^{0000} = +(2^{-5})\ x\ 2^{0} = +2^{-5} \quad \Leftarrow$$

$$\underbrace{0}_{S}\ \underbrace{00011}_{M}\ \underbrace{1001}_{E} \Rightarrow 0\ 0{,}000011\ x\ 2^{1001} = +(2^{-4}+2^{-5})\ x\ 2^{-1} = +(2^{-5}+2^{-6}) \quad \Leftarrow$$

$$\underbrace{1}_{S}\ \underbrace{11111}_{M}\ \underbrace{1111}_{E} \Rightarrow 1\ 0{,}11111\ x\ 2^{1111} = -(2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5})\ x\ 2^{-7} =$$
$$= -(2^{-8}+2^{-9}+2^{-10}+2^{-11}+2^{-12}) \quad \Leftarrow$$

$\underbrace{0}_{S}\,\underbrace{00000}_{M}\,\underbrace{0000}_{E} \Rightarrow 0\ 0{,}00000\ x\ 2^{0000} = +0\ x\ 2^{0} = 0 \quad \Leftarrow$

$\underbrace{0}_{S}\,\underbrace{00000}_{M}\,\underbrace{1111}_{E} \Rightarrow 0\ 0{,}00000\ x\ 2^{1111} = +0\ x\ 2^{-7} = 0 \quad \Leftarrow$

$\underbrace{1}_{S}\,\underbrace{11111}_{M}\,\underbrace{0000}_{E} \Rightarrow 1\ 0{,}11111\ x\ 2^{0000} = -(2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5})\ x\ 2^{0} =$

$= -(2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5}) \quad \Leftarrow$

$\underbrace{1}_{S}\,\underbrace{00000}_{M}\,\underbrace{0000}_{E} \Rightarrow 1\ 0{,}00000\ x\ 2^{0000} = -(0\ x\ 2^{0}) = -0\ x 1 = -0 \Leftarrow$

$\underbrace{0}_{S}\,\underbrace{00001}_{M}\,\underbrace{1111}_{E} \Rightarrow 0\ 0{,}00001\ x\ 2^{1111} = +(2^{-5}\ x\ 2^{-7}) = +2^{-12} \quad \Leftarrow$

$2)\ M \begin{cases} fraccionaria \\ BCS\ 5\ bits \end{cases} \qquad E \begin{cases} BSS \\ 3\ bits \end{cases}$

a) Sin normalizar

$\underbrace{0}_{S}\,\underbrace{1000}_{M}\,\underbrace{111}_{E} \Rightarrow 0\ 0{,}1000\ x\ 2^{111} = +(2^{-1})\ x\ 2^{7} = +(2^{6}) = 64 \quad \Leftarrow$

Normalizada

$\underbrace{0}_{S}\,\underbrace{1000}_{M}\,\underbrace{111}_{E} \Rightarrow 0\ 0{,}1000\ x\ 2^{111} = +(2^{-1})\ x\ 2^{7} = +(2^{6}) = 64 \quad \Leftarrow$

Con bit implícito

$\underbrace{0}_{S}\,\underbrace{1000}_{M}\,\underbrace{111}_{E} \Rightarrow 0\ 0{,}11000\ x\ 2^{111} = +(2^{-1}+2^{-2})\ x\ 2^{7} =$

$= +(2^{6}+2^{5}) = 64+32 = 96 \quad \Leftarrow$

---

b) Sin normalizar

$$\underbrace{1}_{S}\underbrace{1000}_{M}\underbrace{011}_{E} \Rightarrow 1\,0{,}1000\ x\ 2^{011} = -(2^{-1})\ x\ 2^{3} = -(2^{2}) = -4 \quad \Leftarrow$$

Normalizada

$$\underbrace{1}_{S}\underbrace{1000}_{M}\underbrace{011}_{E} \Rightarrow 1\,0{,}1000\ x\ 2^{011} = -(2^{-1})\ x\ 2^{3} = -(2^{2}) = -4 \quad \Leftarrow$$

Con bit implícito

$$\underbrace{1}_{S}\underbrace{1000}_{M}\underbrace{011}_{E} \Rightarrow 1\,0{,}11000\ x\ 2^{011} = -(2^{-1}+2^{-2})\ x\ 2^{3} =$$

$$= -(2^{2}+2^{1}) = -(4+2) - 6 \quad \Leftarrow$$

$$-----------------------------$$

c) Sin normalizar

$$\underbrace{0}_{S}\underbrace{0000}_{M}\underbrace{000}_{E} \Rightarrow 0\,0{,}0000\ x\ 2^{000} = +0 \quad \Leftarrow$$

Normalizada

No se puede. La mantisa no empieza con 0,1...

Con bit implícito

$$\underbrace{0}_{S}\underbrace{0000}_{M}\underbrace{000}_{E} \Rightarrow 0\,0{,}10000\ x\ 2^{000} = +(2^{-1})\ x\ 2^{0} =$$

$$= +(2^{-1})\ x\ 1 = +0{,}5 \quad \Leftarrow$$

$$---------------------------$$

$d)$ *Sin normalizar*

$$\underset{S}{1}\ \underbrace{1111}_{M}\ \underbrace{111}_{E} \Rightarrow 1\,0{,}1111\; x\; 2^{111} = -(2^{-1}+2^{-2}+2^{-3}+2^{-4})\; x\; 2^{7} =$$

$$= -(2^{6}+2^{5}+2^{4}+2^{3}) = -(64+32+16+8) = -120 \quad \Leftarrow$$

Normalizada

$$\underset{S}{1}\ \underbrace{1111}_{M}\ \underbrace{111}_{E} \Rightarrow 1\,0{,}1111\; x\; 2^{111} = -(2^{-1}+2^{-2}+2^{-3}+2^{-4})\; x\; 2^{7} =$$

$$= -(2^{6}+2^{5}+2^{4}+2^{3}) = -(64+32+16+8) = -120 \quad \Leftarrow$$

Con bit implícito

$$\underset{S}{1}\ \underbrace{1111}_{M}\ \underbrace{111}_{E} \Rightarrow 1\,0{,}11111\; x\; 2^{111} = -(2^{-1}+2^{-2}+2^{-3}+2^{-4}+2^{-5})\; x\; 2^{7} =$$

$$= -(2^{6}+2^{5}+2^{4}+2^{3}+2^{2}) = -(64+32+16+8+4) = -124 \quad \Leftarrow$$

- $2^{-9} = \dfrac{1}{2^{9}}$
-

- 0,11001
- 0,00001
- 0,11010

3)a) M BSS frac 8 bits, E BSS 4 bits

- Más grande=N2 = $0,11111111 x 2^{15}$
- $=0,11111111 x 2^{1111} = (2^{-1} + 2^{-2} + 2^{-3} + 2^{-4} + 2^{-5} + 2^{-6} + 2^{-7} + 2^{-8}) x 2^{15} = (1- 2^{-8}) x\, 2^{15}$
- Más chico=N1=$0,00000000 x 2^{E} =0$
- Rango = [N1; N2]=$[0; (1- 2^{-8}) x\, 2^{15}]$

Resolución

N1=0 N4 N3 N2

- N2 = $0,11111111 x 2^{15}$
- N3 = $0,11111110 x 2^{15}$
- Resolución1=N2- N3=$(0,11111111-0,11111110) x 2^{15} =(0,00000001) x 2^{15} = 2^{-8} x 2^{15}$ (peor)= 2^{7}
- N1=$0,00000000 x 2^{E} =0$
- N4=$0,00000001 x 2^{0} =2^{-8}\, 2^{0}$
- Resolución2=N4- N1=$(0,00000001-0) x 2^{0} = 2^{-8}$
- (mejor)

7)P3) 8,625

- Primero la parte entera= 1000=8
- Parte fraccionaria
- 0,625x2=1,25 → me quedo con el 1->0,25
- 0,25x2=0,5 → me quedo con 0 ->0,25
- 0,5x2=1,0 →me quedo con el 1 ->0,0
- $1000{,}10100x2^0$ = $0{,}100010100x2^4$ =8,625
- → 5 bits M → $0{,}10001x2^4$ = $(2^{-1} + 2^{-5}) \times 2^4$ = = 8+ 0,5= 8,5

El que le sigue a 8,5

- $0{,}10001x2^4$ $+0{,}00001x2^4$ = $0{,}10010x2^4$
- El que le sigue a 8,5 es = $(2^{-1} + 2^{-4})\, x2^4$
- (0,5+0,0625)x16=9
- E1= 8,625 – 8,5 =0,125 (error absoluto)
- E2=9 - 8,625 =0,375 (error absoluto)
- Como E1 es el error más chico → 8,5 es la representación más cercana a 8,625
- Error relativo= error absoluto/8,625=0,125/8,625

- 10100000000000000000000
- M=1,10100000000000000000000=
- 1+0,5+0,125=1,625