Azmani Sultana
Id : 22201949
Section: 17

CSE330
Assignment 1

Ans no 1

a) $\beta = 2, m = 4, -2 \le e \le 6$

Convention 1 / lecture note form / General form:

$$F = \pm (0.d_1 d_2 d_3 d_4)_\beta \times \beta^e$$

$$= \pm (0.1 d_2 d_3 d_4)_2 \times 2^9$$

Maximum number $= (0.1111)_2 \times 2^6$

$$= 0.9375 \times 2^6$$

$$= 60$$

Minimum number $= -(0.1111)_2 \times 2^6$

$$= -60$$

Normalized form:

$$F = \pm (1.d_1 d_2 d_3 d_4)_\beta \times \beta^e$$

$$= \pm (1.d_1 d_2 d_3 d_4)_2 \times 2^9$$

Maximum number $= (1.1111)_2 \times 2^6$

$$= 1.9375 \times 2^6$$

$$= 124$$

Minimum number $= -(1.1111)_2 \times 2^6$

$$= -124$$

Denormalized form:

$$F = \pm (0.1d_1d_2d_3d_4)_2 \times 2^3$$

Maximum number, $F = (0.11111)_2 \times 2^6$

$= 0.96875 \times 2^6$

$= 62$

Minimum number, $F = -(0.11111)_2 \times 2^6$

$= -62$

b) General form:

Non-negative minimum number $= (0.1000)_2 \times 2^{-2}$

$= 0.5 \times 2^{-2}$

$= 0.125$

Normalized form:

Non-negative minimum number $= (1.0000)_2 \times 2^{-2}$

$= 1 \times 2^{-2}$

$= 0.25$

Denormalized form:

Non-negative minimum number $= (0.10000)_2 \times 2^{-2}$

$= 0.5 \times 2^{-2}$

$= 0.125$

c) $-2 \le e \le 6$

$e = [-2, 6]$

Total exponent $= 9$

General form:

$$\text{Total number} = 2^3 \times 9 \times 2$$

$$= 144$$

Normalized form:

$$\text{Total number} = 2^4 \times 9 \times 2$$

$$= 288$$

Denormalized form:

$$\text{Total number} = 2^4 \times 9 \times 2$$

$$= 288$$

Ans no 2

a) $x = (5.625)_{10}$

2 | 5	
2 | 2 - 1	↑
2 | 1 - 0	
0 - 1	

$(5)_{10} = (101)_2$

$0.625 \times 2 = 1.25$

$0.25 \times 2 = 0.5$

$0.5 \times 2 = 1.00$

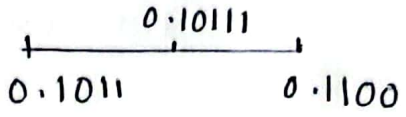
$(0.625)_{10} = (.101)_2$

$\therefore (5.625)_{10} = (101.101)_2$

b) $x = (101.101)_2$

$= (0.101101)_2 \times 2^3$

$fl(x) = (0.1011)_2 \times 2^3$



~~$= 0.6875 \times 2^3$~~

~~$= 5.5$~~

c) $fl(x) = (0.1011)_2 \times 2^3$

$= 0.6875 \times 2^3 = 5.5$

Rounding error, $\delta = \dfrac{|fl(x) - x|}{|x|}$

$= \dfrac{|5.5 - 5.625|}{|5.625|}$

$= 0.022222$

Machine epsilon, $\epsilon = \frac{1}{2}\beta^{-m}$

~~$= \frac{1}{2} \times 2^{-3}$~~

~~$= 0.0625$~~

$= \frac{1}{2} \times 10^{-3}$

$= 5 \times 10^{-4}$

Ans no 3

a) $x^2 - 60x + 1 = 0$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$x_1 = \frac{60 + \sqrt{(-60)^2 - 4.1.1}}{2.1} = 30 + \sqrt{899}$$

$$= 31 + 29.983287$$

$$= 59.983287$$

$$x_1 = 30 + 29.9833$$

$$= 59.9833 \quad (6 \text{ significant figure})$$

$$x_2 = \frac{60 - \sqrt{(-60)^2 - 4.1.1}}{2.1}$$

$$= 30 - \sqrt{899}$$

$$= 30 - 29.983287$$

$$= 0.016712988 7$$

$$x_2 = 30 - 29.9833$$

$$= 0.0167000 \quad (6 \text{ significant figure})$$

$$\delta_1 = \frac{|fl(x_1) - x_1|}{|x_1|}$$

$$= \frac{|59.9833 - 59.983287|}{|59.983287|}$$

$$= 4.784662776 \times 10^{-7} \times 100$$

$$= 0.0000478466277 6 \%$$

$$\delta_2 = \frac{|\delta_1(x_2) - x_2|}{|x_2|}$$

$$= \frac{|0.0167 - 0.0166712987|}{|0.0166712987|}$$

$$= 1.721589315 \times 10^{-3} \times 100$$

$$= 0.1721589315\ \%.$$

$\delta_2 > \delta_1$

So, the loss of significance problem occurs when I calculate the second root.

b) $x_1 = 59.9833$

$x_2 = 0.0167000$

$$x_1 x_2 = \frac{c}{a}$$

$$\Rightarrow x_2 = \frac{c}{a x_1}$$

$$= \frac{1}{1 \times 59.9833}$$

$$= 0.0166713068 5$$

$$= 0.0166713 \quad (6 \text{ significant figure})$$