ASSIGNMENT-1

→ Sangram Patro

Q3) a, b, c, d → programme is attached.

Q3) e) For 48-bit arithmetic

Sum will be between 16 − 32. So the process will stop at $i = 2^{44} \simeq 1.76 \times 10^{13}$

For 50-bit arithemetic

Summation will stop when it becomes 32.

Q6)

In iii) & (iv) $n = 2^i$ so we can compare the effective rate of convergence of ii, iii, iv.

And (iii) & (iv) are less efficient as far as (ii) is concerned as it requires 2 square roots at each step.

while calculating the value of $\sin\theta$, we subtract nearly equal numbers (in (iii) & (iv)) so it is unstable. So (ii) can achieve full accuracy as compared to other 3.

7. (i) $101101.101_2 \times 1010.011_2$

$$
\begin{array}{r}
101101101 \\
1011011010 \\
1011011010000 \\
10110110100000 \\
\hline
(111011001.010111)_2
\end{array}
$$

Verify

$$45.625_{10} \times 10.375_{10}$$

$$473.359375_{10}$$

$473.359375_{10} = (111011001.010111)_2$

(ii) $571.45_8 \times 65.35_8$

$$
\begin{array}{r}
353771 \\
2154570 \\
35377100 \\
433136000 \\
\hline
(47326.5661)_8
\end{array}
$$

Verify

$$377.578125_{10} \times 53.453125_{10}$$

$$(20182.730712890625)_{10}$$

$= 4\times8^4 + 7\times8^3 + \dots\dots + 1\times8^{-4} = 20182.730712890625$

(iii) $51.A_{16} \times B.3_{16}$

$$
\begin{array}{r}
F4E \\
381E0 \\
\hline
391.2E_{16}
\end{array}
$$

Verify

$$11.1875_{10} \times 81.625_{10}$$

$$(913.1796875)_{10}$$

$391.2E_{16} = 3\times16^2 + 9\times16^1 + \dots\dots + (14\times16^{-2})$

$= 913.1796875_{10}$

(iv) $\bar{1}10\bar{1}0.0\bar{1}_3 \times \bar{1}10.1\bar{1}_3$ (balanced ternary)

$$
\begin{array}{r}
1\bar{1}01001 \\
\bar{1}10\bar{1}00\bar{1}0 \\
1\bar{1}10\bar{1}00\bar{1}000 \\
1\bar{1}0100 10000 \\
\hline
\cancel{000\bar{1}1\bar{1}000\bar{1}1} \\
01101\bar{1}000\bar{1}1_3
\end{array}
$$

Verify

$$-57.\bar{1}_{10} \times -5.\bar{7}_{10}$$

$$329.975308_{10}$$

7 b) (i) Binary

- 23.35

$23 + 0.35 = 23 + 0.35$

$= 10111_2 + 0.01\overline{0110}_2$

$= 10111.01\overline{0110}_2$

- −23.35

$= -(010111.01\overline{01110})_2 = 101000.101001_2$

$+ 0.00001_2$

$= 101000.101010_2$

(iii) Negative decimal

23.35

- $23.35 = 183.75_{-10}$
- $-23.35 = 37.45_{-10}$
- $\frac{22}{7} = 3.\overline{958}_{-10}$

(iv) Balanced ternary

- $23.35 = 10\bar{1}\bar{1}.\overline{1001}$
- $-23.35 = -(10\bar{1}\bar{1}.\overline{1001})$

$\cancel{\frac{22}{7}}$ $= \bar{1}011.\overline{\bar{1}00\bar{1}}$

- $\frac{22}{7} \simeq 3.142857143 = 3.\overline{142857}$

$= 10.\overline{0110\bar{1}\bar{1}}$

7C (i) $53 + 45 = 98$

$00110101 = 53, 00101101 = 45$

$00101101 = 45$

$01100010 = 98$ ✓

(ii)

$01010011 = 83$

$01011100 = 92$

$10101111 = 175 - 256$

$= -81$ ✓

(iii) $-53 + (-45) = -98$

$11001011 \rightarrow -53$

$11010011 \rightarrow -45$

$10011110 \rightarrow -98$ ✓

(iv) $52 + (-73) = -21$

$00110100 = 52$

$10110111 = -73$

$11101011 = -21$ ✓

(v) $-53 + (-75) = -128$

$11001011 = -53$

$10110101 = -75$

~~01010001~~

~~1111110~~

$10000000 = -128$ ✓

(vi) $-83 + (-92) = -175$

$10101101 = -83$

$10100100 = -92$

$01010001 = -175 + 256$

$= 81$ ✓

(vii) $53 + 75 = 128$

$00110101 = 53$

$01001011 = 75$

$10000000 = -128$

$= +128 - 256$ ✓

(viii) $73 + (-73) = 0$

$01001001 \rightarrow 73$

$10110111 \rightarrow -73$

$00000000 = 0$ ✓

(8) a) -3.14

$-3.14 = 11.00100001111010111000001_2$

$= 1.100100001111010111000001 \times 10^1$

-3.14 In IEEE floating point representation for 32-bit

number is :- $1 \quad \underbrace{10000000}_{128_{10}} \quad 10010000111101011100001$

$128 = 2^{8-1}$

(b) $\frac{22}{7} \approx 3.\overline{142857}$

$= 11.\overline{001} = 1.1\overline{001} \times 10^1$

In IEEE floating point representation :

$0 \quad \underbrace{10000000}_{} \quad 10010010010010010010010$

$= 128(10)$

$= 2^{8-1}$