Yitian Huang
2/5/2022

CS 3010 S22 Assignment 1

1 | 8 | 23

1a $2^{-30}_{10} = 0.000\ldots001_2 = 1.0 \times 2^{-30}$ (1.0 → mantissa)

$-30 + 127 = 97$

97	1
48	0
24	0
12	0
6	0
3	1
1	1

$97_{10} = 0110\ 0001_2$ → exponent

$2^{-30} = 0\ 0110\ 0001\ 0000\ 0000\ 0000\ 0000\ 0000\ 000$

1b $64.015625_{10} = 1000000.000001_2 = 1.000000000001 \times 2^6$ (mantissa)

64	0	0.015625
32	0	0.03125 | 0
16	0	0.0625 | 0
8	0	0.125 | 0
4	0	0.25 | 0
2	0	0.5 | 0
1	1	1 | 1

$6 + 127 = 133$

133	1
66	0
33	1
16	0
8	0
4	0
2	0
1	1

$133_{10} = 1000\ 0101_2$ → exponent

$64.015625 = 0\ 1000\ 0101\ 0000\ 0000\ 0001\ 0000\ 0000\ 000$

1c $-8 \times 2^{-24} = -(2^3 \times 2^{-24}) = -2^{-21} = -0.00\ldots01_2 = 1.0 \times 10^{-21}$ (1.0 → mantissa)

$-21 + 127 = 106$

106	0
53	1
26	0
13	1
6	0
3	1
1	1

$106_{10} = 0110\ 1010_2$ → exponent

$-8 \times 2^{-24} = 2^{-4} = 1\ 0110\ 1010\ 0000\ 0000\ 0000\ 0000\ 0000\ 000$

1d $0.5_{10} = 0.1_2 = 1.0 \times 2^{-1}$ (1.0 → mantissa)

$-1 + 127 = 126$ $\quad 126_{10} = 0111\ 1110_2$ exponent

0.5	
1	1

126	0
63	1
31	1
15	1
7	1
3	1
1	1

$0.5 = 0\ 0111\ 1110\ 0000\ 0000\ 0000\ 0000\ 0000\ 000$

$5 + 127 = 132$

1e $42.424286 = 101010.011011001001101100_2 = 0\ 1000\ 0100\ 0101\ 0011\ 0110\ 0100\ 1101\ 100$

42	0	.	0	1	1	0	1	1	0	0	1	0	0	1	1	0	1	1	0	0		
21	1		0.424286	0.848572	1.697144	1.394288	0.788576	1.577152	1.154304	0.308608	0.617216	1.234432	0.468864	0.937728	1.875456	1.750912	1.501824	1.003648	0.007296	0.014592	0.029184	0.058368
10	0																					
5	1																					
2	0																					
1	1																					

1f $76.234567 \times 10^{-15} = 1.6101011101010100011100 \times 2^{-44}$

$-44 + 127 = 83 = 01010011$ $\quad 76.234567 \times 10^{-15} = 0\ 0101\ 0011\ 1010\ 1101\ 0101\ 0001\ 1100$

1g $1.4345678 = 1.0110\ 1111\ 0011\ 1111\ 1010\ 10 = 0011\ 1111\ 0110\ 1111\ 0011\ 1111\ 1010\ 1011$

0.43	0	.	0	1	1	0	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	0	1	0
0.86	1		0.72	1.44	0.88	1.76	1.52	1.04	0.08	0.16	0.32	0.64	1.28	0.56	1.12	0.24	0.48	0.96	1.92	1.84	1.68	1.36	0.72	1.44
1.72	1																							
1.44	1																							
0.88	0																							
1.76	1																							
1.52	1																							
1.04	1																							
0.08	0																							
0.16	0																							
0.32	0																							
0.64	0																							
1.28	1																							
0.56	0																							
1.12	1																							
0.24	0																							
0.48	0																							
0.96	0																							
1.92	1																							
1.84	1																							
1.68	1																							
1.36	1																							
0.72	0																							
1.44	1																							

$0 + 127 = 127$

01111111_2 → exponent.

1h $3.14159265358979323846264338327950288_4 = 11.001\ 0010\ 0000\ 1111\ 1101\ 1011$

3	1
1	1

$1 + 127 = 128 = 10000000_2$ → exponent $= 0\ 1000\ 0000\ 1001\ 0010\ 0001\ 1111\ 1011\ 011$

1i $3.14_{10} = 11.0010\ 0011\ 1101\ 0111\ 0000\ 10_2 = 0\ 1000\ 0000\ 1001\ 0001\ 1110\ 1011\ 1000\ 01$

1	1	0 0 1 0 0 0 1 1 1 1 0 1 0 1 1 1 0 0 0 0 1 0
1	1	128 64 32 16 8 4 2 1 0.5 0.25 0.125 0.0625 0.03125 0.015625 0.0078125 0.00390625 0.001953125 0.0009765625

$1+127=128 = 1000\ 0000$ → exponent.

1j $17/31 \approx 0.54838709677419354839 = 0.10001100\ 0110\ 0011\ 0001\ 1001$

$-1 + 127 = 126 = 0111\ 1110$ → exponent

1 0 0 0 1

$17/31 = 0\ 0111\ 1110\ 0001\ 1000\ 1100\ 0110\ 0011\ 0011$

2a $0\ 0000\ 0000\ 0000\ 0000\ 0000\ 0000\ 0000 = 0.0$ because all 0 is reserved

2b. $1\ 1111\ 1111\ 0000\ 0000\ 0000\ 0000\ 0000\ 000 = -\inf$ because its also reserved.

2c $0\ 1000\ 0001\ 0110\ 0000\ 0000\ 0000\ 0000\ 000$ $1.0110_2 = 1.375_{10}$

$1000\ 0001_2 = 129_{10}$ $129-127=2$ $2^0+2^{-2}+2^{-3}=1.375$ 1.375×2^2 or 5.5

2d $0\ 01111111\ 0\ldots0$ $1.0_2 = 1_{10}$ $1\times 2^0 = 1$

$01111111_2 = 127_{10}$ $127-127=0$.

2e $0\ 0111\ 1011\ 1001\ 1001\ 1001\ 1001\ 1001\ 100$ $1.1001\ 1001\ 1001\ 1001\ 1001\ 100 \approx 1.5999999046326$

(-1 -4 -5 -8 -9 -12 -13 -16 -17 -20 -21)

$01111011_2 = 2^0+2^1+2^3+2^4+2^5+2^6 = 123$ $123-127=-4$ $\approx 1.5999999046326\times 2^{-4} = 0.09999999403953$

Relative Error $= \frac{|\alpha-\beta|}{|\alpha|} = \frac{|(x'-y')-(x-y)|}{|(x-y)|}$

3a $\frac{|1.5673456545535678906211 - 1.567\ 3456\ 5456|}{|1.5673456545556789062 11|} = 2.0486469533117\times 10^{-13}$

3b $y' = 0.00781250000$ $\frac{|(1-0.00781250000)-(1-2^{-7})|}{|(1-2^{-7})|} = 0$ too small

3c $y' = 0.00000000000$ $\frac{|(1-0.000000000000)-(1-10^{-13})|}{|(1-10^{-12})|} = 1.0\times 10^{-13}$

3d $2.3456 - 0.00000456789011234$ $y' = 0.00000456789$

$\frac{|2.3456-0.00000456789)-(2.3456-0.00000456789011234)|}{|(2.3456-0.00000456789011234)|} = 4.78949\times 10^{-14}$

3e $x = 3.14562345678912$ $y = 3.14562345123456$. relative error $= \frac{|(x'-y')-(x-y)|}{|(x-y)|}$

$x' = 3.1456234567891$ $y' = 3.1456234512345$

$= 6.4078979826264\times 10^{-5}$