Assignment_4

3.23
Write down the binary representation of the decimal number 63.25 assuming the IEEE 754 single precision format.

$63.25 = 63 + 0.25$

$63 = 0x3f = 0b00111111 = 111111_2$

$0.25 = 0.25 \times 4 / 4 = 1/4 = 1 / 2^2 = 1 \times 2^{-2} = 0.01_2$

$63.25 = 63 + 0.25 = 111111_2 + 0.01_2 = 111111.01_2 = 1.1111101_2 \times 2^5$

Another approach*:

$0.25 = 0.5 \times 2^{-1}$

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

$= 0.01_2$

Sign bit = 0

Fraction = 1111 1010 0000 0000 0000 000

Exponent – Bias = 5

Exponent = 5 + Bias = 5 + 127 = 132 = 0x84 = 10000100_2

Final bit pattern: 0 1000 0100 1111 1010 0000 0000 0000 000

3.24 Write down the binary representation of the decimal number 63.25 assuming the IEEE 754 double precision format.

(only answer is provided)

Final bit pattern:

0 100 0000 0100 1111 1010 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000

3.25 Write down the binary representation of the decimal number 63.25 assuming it was stored using the single precision IBM format (base 16, instead of base 2, with 7 bits of exponent).

The IBM floating point number is represented as the following formula:

$(-1)^{sign} \times 0.\text{fraction}_{16} \times 16^{(\text{exponent}-64)}$

$63.25 = 63 + 0.25 = 111111_2 + 0.01_2 = 111111.01_2 = 111111.01_2 \times 2^0 = 111111.01_2 \times 16^0 =$

$111111.0100_2 \times 16^0 = 3f.4_{16} \times 16^0 = .3f4_{16} \times 16^2$

Sign = 0

Fraction = $3f4_{16}$ =0011 1111 0100…0_2 (24 bits)

Exponent – 64 = 2, thus exponent = 64 + 2 = 66 = 0x42 = 1000010_2 (7 bits)

Final bit pattern: 0 1000010 001111110100000000000000

3.41 Using the IEEE 754 floating point format, write down the bit pattern that would represent -1/4. Can you represent -1/4 exactly?

-1/4 = -0.25

$0.25 = 0.25 \times 4 / 4 = 1/4 = 1 / 2^2 = 1 \times 2^{-2} = 0.01_2 = 1.0 \times 2^{-2}$

Sign = 1 (negative)

Fraction = 0….0 (23 bits)

Exponent – bias = Exponent – 127 = -2, thus, Exponent = 125 = 0x7d = 01111101_2

Final bit pattern (an exact representation)

1 01111101 00000000000000000000000

*