Princess Sumaya University for Technology
Computer Engineering Dept.
22541 - Computer Architecture - Fall 2023
Quiz 3

Name: ____ *Solution* _____   Duration: 5 minutes.
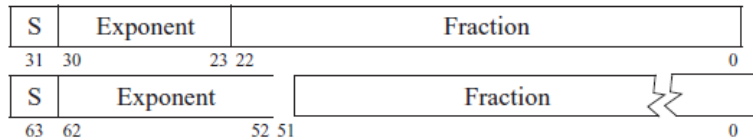
Instructions: - *No questions allowed.*   *-Show your work.*

IEEE 754 FLOATING-POINT STANDARD

$$(-1)^S \times (1 + \text{Fraction}) \times 2^{(\text{Exponent - Bias})}$$

where Single Precision Bias = 127,
Double Precision Bias = 1023.

IEEE Single Precision and Double Precision Formats:

④ **IEEE 754 Symbols**

Exponent	Fraction	Object
0	0	± 0
0	≠0	± Denorm
1 to MAX - 1	anything	± Fl. Pt. Num.
MAX	0	±∞
MAX	≠0	NaN

S.P. MAX = 255, D.P. MAX = 2047

S	Exponent	Fraction
31 30	23 22	0

S	Exponent	Fraction
63 62	52 51	0

1. For the following floating-point number:

1100 0011 0101 1000 0000 0000 0000 0000

a) What is the number precision?   (1pt)

   Single Precision

b) What is the number in normalized binary format?   (3pts)

   1000 0110 = 134

   134-127 = 7

   -1.1011×2^7

2. Write the floating-point binary representation of:

   • +Zero   (2pts)

0	0000 0000	000 0000 0000 0000 0000 0000

   • +infinity   (2pts)

0	1111 1111	000 0000 0000 0000 0000 0000

   • Any negative denormalized number   (2pts)

1	0000 0000	*Any binary combination other than* 000 0000 0000 0000 0000 0000

Name: __ Solution _____   Duration: 5 minutes.
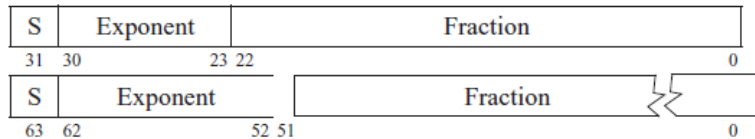
Instructions: - *No questions allowed.*   -*Show your work.*

IEEE 754 FLOATING-POINT STANDARD

$$(-1)^S \times (1 + \text{Fraction}) \times 2^{(\text{Exponent - Bias})}$$

where Single Precision Bias = 127,
Double Precision Bias = 1023.

IEEE Single Precision and Double Precision Formats:

IEEE 754 Symbols ④

Exponent	Fraction	Object
0	0	± 0
0	≠0	± Denorm
1 to MAX - 1	anything	± Fl. Pt. Num.
MAX	0	±∞
MAX	≠0	NaN

S.P. MAX = 255, D.P. MAX = 2047

S	Exponent	Fraction
31 30	23 22	0

S	Exponent	Fraction
63 62	52 51	0

1. For the following floating-point number:

1100 0010 0100 1100 0000 0000 0000 0000

a) What is the number precision?   (1pt)

   Single Precision

b) What is the number in normalized binary format?   (3pts)

1000 0100 = 132

132-127= 5

   -1.10011 x 2⁵

2. Write the floating-point binary representation of:

   • Negative Zero   (2pts)

1	0000 0000	000 0000 0000 0000 0000 0000

   • -infinity   (2pts)

1	1111 1111	000 0000 0000 0000 0000 0000

   • Any negative denormalized number   (2pts)

1	0000 0000	Any binary combination other than 000 0000 0000 0000 0000 0000