

Last NAME: Shameem

First Name: Azwad

Computer Science C.Sc. 342

Quiz No.1

Time of performance 5:00-6:15 PM on February 23, 2022 Please

write your Last Name on every page:

NO CORRECTIONS ARE ALLOWED IN ANSWER CELLS!!!!

You may use the back page for computations.

Please answer all questions. **Not all questions are of equal difficulty. Please review the entire quiz first and then budget your time carefully.**

Please **HAND WRITE** and sign statements affirming that you will not cheat:

"I will neither give nor receive unauthorized assistance on this exam. I will use only one computing device to perform this test"

Please **HAND WRITE** and sign here:

I will neither give nor receive unauthorized assistance on this exam. I will only use one computing device to perform this test
Azwad Shameem

1. [10 points] For **each 8 BIT** binary pattern shown in the table below please write corresponding values of the following interpretations: **UNSIGNED INT**, **SIGNED INT**, **UNSIGNED Fixed Point**, **SIGNED Fixed Point**.

Each correctly answered column is **2.5** points. **FIXED POINT IS LOCATED TWO POSITIONS FROM THE RIGHT!**

MOST SIGNIFICANT BIT IS 7. LEAST SIGNIFICANT BIT IS 0.

76543210	UNSIGNED INT	SIGNED INT	UNSIGNED Fixed Point	SIGNED Fixed Point
10000000	$2^7 = 128$	$-(2^7) = -128$	$2^5 = 32$	$-(2^5) = -32$
10000011	$2^7 + 2^1 + 2^0 = 131$	$-(2^7) + 2^1 + 2^0 = -125$	$2^5 + 1/2 + 1/2^2 = 131/4$	$-(2^5) + 1/2 + 1/4 = -125/4$
10000001	$2^7 + 2^0 = 129$	$-(2^7) + 2^0 = -127$	$2^5 + 1/4 = 129/4$	$-(2^5) + 1/4 = -127/4$
01000001	$2^6 + 2^0 = 65$	$2^6 + 2^0 = 65$	$2^4 + 1/4 = 65/4$	$2^4 + 1/4 = 65/4$
01111111	$2^7 - 1 = 127$	$2^7 - 1 = 127$	$2^4 + 2^3 + 2^2 + 2^1 + 2^0 + 1/2 + 1/4 = 127/4$	$2^4 + 2^3 + 2^2 + 2^1 + 2^0 + 1/2 + 1/4 = 127/4$
11111111	$2^7 + 2^6 + 2^5 + 2^4 + 2^3 + 2^2 + 2^1 + 2^0 = 255$	$-(2^7) + 2^6 + 2^5 + 2^4 + 2^3 + 2^2 + 2^1 + 2^0 = -1$	$2^5 + 2^4 + 2^3 + 2^2 + 2^1 + 2^0 + 1/2 + 1/4 = 255/4$	$-(2^5) + 2^4 + 2^3 + 2^2 + 2^1 + 2^0 + 1/2 + 1/4 = -1/4$
11111100	$2^7 + 2^6 + 2^5 + 2^4 + 2^3 + 2^2 = 252$	$-(2^7) + 2^6 + 2^5 + 2^4 + 2^3 + 2^2 = -4$	$2^5 + 2^4 + 2^3 + 2^2 + 2^1 + 2^0 = 63$	$-(2^5) + 2^4 + 2^3 + 2^2 + 2^1 + 2^0 = -1$
00000000	0	0	0	0
01111110	$2^6 + 2^5 + 2^4 + 2^3 + 2^2 + 2^1 = 126$	$2^6 + 2^5 + 2^4 + 2^3 + 2^2 + 2^1 = 126$	$2^4 + 2^3 + 2^2 + 2^1 + 2^0 + 1/2 = 63/2$	$2^4 + 2^3 + 2^2 + 2^1 + 2^0 + 1/2 = 63/2$
10001110	$2^7 + 2^3 + 2^2 + 2^1 = 142$	$-(2^7) + 2^3 + 2^2 + 2^1 = -114$	$2^5 + 2^1 + 2^0 + 1/2 = 71/2$	$-(2^5) + 2^1 + 2^0 + 1/2 = -57/2$
00010011	19	+19	$\frac{3}{4+4} = \frac{16+3}{4}$	$\frac{3}{4+4} = \frac{19}{4}$

Fixed Point

2. [10 points] What is the most negative number (largest absolute value negative) that can be represented using 16 bit signed integer representation? Please CIRCLE AROUND over all the correct ones:

-32768, -65536, -16384, -32767, NONE

3. [10 points] Please subtract two number in Hex. Then convert each operand to binary and perform the same operation in binary, then repeat BASE 10. The signed integers are represented using two's complement.

0x0E	00001110	14 (14 * 160)
- (+: two negatives make positive)	-	-
0xFF (0x00 + 0x01 = 0x01) (Signed complement)	11111111 (00000000 + 1 = 00000001)	-1 (0x00 + 1 = 0x01)

Result: 0x0F

0000 0000b: 00001111

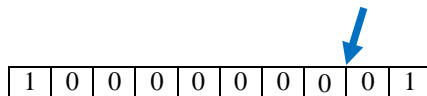
dec: +15

4. [20 points]

Determine the **MINIMAL** number of **bits** required to represent **-127.75** using:

4.1 (5 points) ASCII code 8 * 7 = 56 bits (please write the number of bits in the cell)

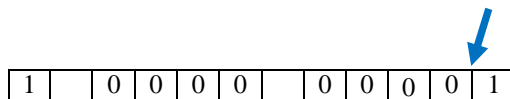
4.2 (5 points) Binary Fixed Point representation 10 bits (please write the number of bits in the cell) And the corresponding binary Fixed Point representation here.



$$(2^7) + 2^{-2} = -128 + .25 = -127.75$$

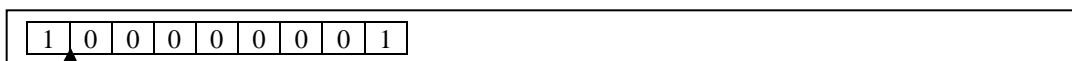
Fixed Point

- 4.3 (5 points) Take the result from you answer in 4.2 and shift fixed point by 2 positions to the **RIGHT**. Please write down the resulting signed decimal value, And the corresponding binary Fixed Point representation here.



$$-(2^9) + 2^0 = -512 + 1 = -511$$

4. 4(5 points) Please write down the signed rational number stored in the 9-bit word below:



Fixed Point

$$-(2^0) + 2^{-8} = -1 + 1/256 = -255/256$$

5. [10 points] Please determine if single precision floating point representation given below is *NAN, or +Infinity, Infinity, or a valid number* floating point: The top row shows the bit index. **PLEASE JUSTIFY your ANSWER and SHOW your work!** Just the final result will not count as a correct answer.

3	3	2	2	2	2	3	2	2	2	2	2	1	1	1	1	1	1	1	1	1	9	8	7	6	5	4	3	2	1
1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0								
0	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

First bit is 0, therefore it is positive.

Next 8 bits are the exponent, which is $2^7 + 2^0 = 129$.

The final 32 bits show $2^{-1} + 2^{-2} = \frac{3}{4}$.

Bias is 127.

We can now use the equation $(-1)^s * (1 + \text{binary fraction}) * 2^{(\text{exponent} - \text{bias})}$.

$$(-1)^0 * (1 + \frac{3}{4}) * 2^{(129-127)} = (1 + \frac{3}{4}) * 2^2 = \frac{7}{4} * 4 = +7$$

The number +7 is a valid number and not NAN, infinity, or zero.

6. [10 points] Please determine the decimal value (scientific notation) of the single precision floating point representation given below: The top row shows the bit index. **PLEASE SHOW your work!** Just the final result will not count as correct answer. *If it represents NAN, or Infinity, or zero please state this and justify.*

3	3	2	2	2	2	3	2	2	2	2	2	1	1	1	1	1	1	1	1	1	9	8	7	6	5	4	3	2	1
1	0	9	8	7	6	5	4	3	2	1	0	9	8	7	6	5	4	3	2	1	0								
1	1	0	0	0	0	1	0	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0

First bit is 1, therefore it is negative.

Next 8 bits are the exponent, which is $2^7 + 2^2 + 2^0 = 133$.

The final 32 bits show .996039375.

Bias is 127.

We can now use the equation $(-1)^s * (1 + \text{binary fraction}) * 2^{(\text{exponent} - \text{bias})}$.

$$(-1)^1 * (1 + .11111111) * 2^{(133-127)} = -(1.99609375) * 2^{(133-127)} = -1.99609375 * 2^6 = -127.75$$

The number -1.2775×10^2 is a valid number and not NAN, infinity or zero.

7. [5 points] Please determine the decimal value (scientific notation) of the single precision floating point representation given below: The top row shows the bit index. **PLEASE SHOW your** work! Just the final result will not count as correct answer. **If it represents NAN, or Infinity, or zero please state this and justify.**

31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

First bit is 0, therefore it is positive.

Next 8 bits are the exponent, which is just 0.

The final 32 bits show 0.

Bias is 127.

This is a special case because all the exponent values are 0, the invisible leading bit 1 of the mantissa is no longer used, so we have a different formula which is.

$$(-1)^s * (\text{binary fraction}) * 2^{(\text{exponent}-\text{bias})}$$

$$(-1)^0 * (0) * 2^{(0-127)} = +0 * 2^{-127} = +0.$$

The answer is +0, which is a positive zero.

8. [5 points] Please determine the decimal value (scientific notation) of the single precision floating point representation given below: The top row shows the bit index. **PLEASE SHOW your** work! Just the final result will not count as correct answer. If it represents NAN, or Infinity, or zero please state this and justify.

31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

First bit is 1, therefore it is negative.

Next 8 bits are the exponent, which is 11111111 = 255.

The final 32 bits show .111111111111111111111111111111.

Bias is 127.

We can now use the equation $(-1)^s * (1+\text{binary fraction}) * 2^{(\text{exponent}-\text{bias})}$.

This is a special case because all the exponent values are equal to 1 which means it can be either infinity or NAN.

However, since the mantissa is not all 0, we know that it cannot be infinity and therefore the answer must be NAN.

Please write decimal, and binary operands and the results. For each question you have to write the result and **overflow** or **No overflow**. You may override '0' with '1'.

0x0000000E	NO OVERFLOW	14	0000 0000 0000 0000 0000 0000 0000 1101
+ (-): new arithmetic		+	+ (-)
0xFFFFFFFF		(-1)	0000 0000 0000 0000 0000 0000 0000 0001

HEX: 0x0000000D Decimal: $2^3 + 2^2 + 2^0 = 13$ Binary: 0000 0000 0000 0000 0000 0000 0000 1101

0x7FFFFFFF	<div style="border: 1px solid black; padding: 2px; display: inline-block;">OVERFLOW</div>	$2^{31} - 1$	0111 1111 1111 1111 1111 1111 1111 1111
- (+)		- (+)	- (+)
0xFFFFFFFF		1	0000 0000 0000 0000 0000 0000 0000 0001

000000 + 0x00000001 = 0x00000001)

HEX: 0x8000000000 Decimal: $2^{31} = 2147483648$ Binary: 1000 0000 0000 0000 0000 0000 0000 0000

0x80000000	NO OVERFLOW	$-2^{31} - 1$	1000 0000 0000 0000 0000 0000 0000 0000
- (+)		- (+)	- (+)
0xFFFFFFFF		1	0000 0000 0000 0000 0000 0000 0000 0001
000000 + 0x00000001 = 0x00000001)			

HEX: 0x80000001 Decimal: $-2^{31} + 1 = -2147483647$ Binary: 10000000000000000000000000000001

0x7FFFFFFF	<div style="border: 1px solid black; padding: 5px; display: inline-block;">NO OVERFLOW</div>	$2^{31} - 1$	0111 1111 1111 1111 1111 1111 1111 1111
+ (-)		+ (-)	+ (-)
0xFFFFFFFF		1	0000 0000 0000 0000 0000 0000 0000 0001
000000 + 0x00000001 = 0x00000001)			

HEX: 0x7FFFFFFE Decimal: $2^{31} - 2 = 2147483646$ Binary: 0111 1111 1111 1111 1111 1111 1111 1110

Diagram illustrating a 32-bit addition:

```

    0x80000000
  + 0xFFFFFFFF
  -----
    0x7FFFFFFF
  
```

OVERFLOW

Decimal: $+2^{31}-1$ Binary: **01111111111111111111111111111111**