Question 6 [16 points]

Consider the following 8-bit floating point representation based on the IEEE floating point format:

- There is a sign bit in the most significant bit.
- The next 3 bits are the exponent. The exponent bias is $2^{3-1} - 1 = 3$.
- The last 4 bits are the fraction.
- The representation encodes numbers of the form: $V = (-1)^s \times M \times 2^E$, where M is the significand and E is the exponent.

The rules are like those in the IEEE standard (i.e. normalized and denormalized numbers, and the same representation of 0, infinity, and NAN).

Fill in the table below for this format. Here are the instructions for each field:

- **Binary:** The 8 bit binary representation.
- **M:** The value of the significand. This should be a number of the form x or $\frac{x}{y}$, where x is an integer, and y is an integral power of 2. Examples include 0, $\frac{3}{4}$.
- **E:** The integer value of the exponent.
- **Value:** The numeric value represented by the number.

  Note: you need not fill in entries marked with "—".

Description	Binary	M	E	Value
Minus zero				-0.0
—	01000101			
Smallest denormalized				
Largest normalized				
One				1.0
—				5.5
Positive infinity		—	—	$+\infty$

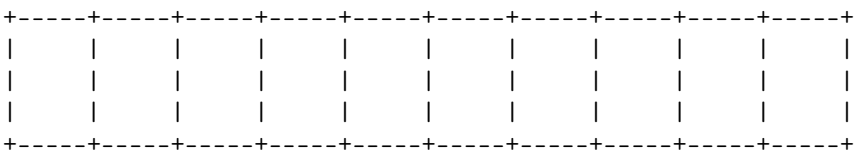
Name: _____ Leginr: _____

Question 11 [16 points]

Consider a floating point format which uses 10 bits but otherwise follows IEEE standard format. 5 bits are used for the fractional part, and 4 bits to represent the exponent.

Sketch the format of this number as bits, with the most significant bit on the left. Mark each bit as **S** for sign, **M** for mantissa, or **E** for exponent.

(2 points)



The **bias** for this format is 7. Explain why.

(2 points)

How is the real number 1/2 represented in binary in this system? Show your working.

(4 points)

[Question continues on the next page]

[continued]

What real number is represented by the binary value 1111000000 in this system? Show your working

(4 points)

What number in this system is represented by the smallest negative denormalized value?

Give your answer as a decimal number, and show your working.

(4 points)

Question 3 [20 points]

Consider a floating point format which uses 9 bits but otherwise follows IEEE standard format. 4 bits are used for the fractional part, and 4 bits to represent the exponent.

What is the *bias* for this format?

(2 points)

What integer is the largest positive normalized value below infinity?

Give your answer as a decimal number, and show your working.

(4 points)

What real number is represented by the binary value 100000000 in this system? Show your working

(2 points)

[Question continues on the next page]

[continued]

How is the real number 1 represented in binary in this system? Show your working.

(4 points)

What real number is represented by the binary value 111100000 in this system? Show your working

(4 points)

[Question continues on the next page]

[continued]

How is the real number -1/32 represented in binary in this system? Show your working.

(4 points)