Name: ________________________________

Below is a toy model of IEEE double-precision floating-point format to demonstrate how machine arithmetic is done.
A number is represented by a sequence of 8 bits:

\[
\begin{align*}
\hat{c}_0 \quad \hat{c}_1 \quad \hat{c}_2 \quad \hat{c}_3 \quad \hat{c}_4 \quad a_1 \quad a_2 \quad a_3
\end{align*}
\]

This sequence represents the number \( x = \sigma \cdot \bar{x} \cdot 2^e \) where

- If \( 1 \leq E \leq 14 \) then
  \[
  \sigma = \begin{cases} 
  1 & \text{if } \hat{c}_0 = 0, \\
  -1 & \text{if } \hat{c}_0 = 1, 
  \end{cases} \\
  e = E - 7, \\
  \bar{x} = (1.a_1a_2a_3)_2
  \]

- If \( E = 0 \) then \( e = -6 \) and \( \bar{x} = (0.a_1a_2a_3)_2 \).

- If \( E = 15 \) then the bit sequence represents \( \pm \infty \) (depending on the sign \( \sigma \)).

(a) Convert the number 28.375 from decimal system to binary system (exact form).

\[28.375 = 28 + 0.375\]

\[28 = (11100)_2\]

\[0.375 = (0.011)_2\]

\[28.375 = (11100.011)_2\]

(b) What is the (approximate) representation of 28.375 in this format?

\[(11100.011)_2 = (1.1100011)_2 \times 2^4 \approx (1.110)_2 \times 2^4\]

(c) Perform the below floating-point multiplication by following the procedure:

1. Add two exponents.
2. Multiply the significands.
3. Normalize the result by shifting the floating point.
4. Round the significand.

\[(1.101)_2 \times 2^{-3} \times (1.111)_2 \times 2^2\]

\[\text{Note that } M = 7 \text{ and } m = -6\]
1) Add exponents:

\[ 2^{-3} \times 2^{2} = 2^{-1} \]

-1 is between m and M.

2) Multiply the significands:

\[
\begin{array}{c}
1.101 \\
\times 1.111 \\
\hline
1.101 \\
1.101 \\
1.101 \\
1.101 \\
\hline
1.000111
\end{array}
\]

1+1 = (10) \_2. Write 0, carry 1.

1+1+1+1 = 4 = (100) \_2. Write 0, carry 10

1+1+10 = (100) \_2. Write 0, carry 10

1+10 = (11) \_2. Write 11

3) Shift the floating point:

\[ xy = (11.000011) \_2 \times 2^{-1} \]

\[ = (1.10100011) \_2 \times 2^{0} \]

4) Round off:

\[ xy \approx (1.100) \_2 \times 2^{0} \]