Math 781 Hw2

due Wednesday 09/07/2022.

1. Let $x = (1.11\cdots111000\cdots)_2 \times 2^{16}$, in which the fraction part has 26 1's followed by 0's. For the Marc-32, determine x_-, x_+, $fl(x)$, $x - x_-$, $x_+ - x$, $x_+ - x_-$, and $\left|\frac{x-fl(x)}{x}\right|$.

2. Which if these is not necessarily true on the Marc-32? (Here x, y, and z are machine numbers and $|\delta| \leq 2^{-24}$.

   (a) $fl(xy) = xy(1+\delta)$

   (b) $fl(x+y) = (x+y)(1+\delta)$

   (c) $fl(xy) = \frac{xy}{1+\delta}$

   (d) $|fl(xy) - xy)| \leq |xy|2^{-24}$

   (e) $fl(x+y+z) = (x+y+z)(1+\delta)$

3. Are these machine numbers in the Marc-32?

   (a) 10^{40}

   (b) $2^{-1} + 2^{-26}$

   (c) $\frac{1}{5}$

   (d) $\frac{1}{3}$

   (e) $\frac{1}{256}$

4. Let $x = 2^{16} + 2^{-8} + 2^{-9} + 2^{-10}$. What is $|x - fl(x)|$ in the Marc-32?

5. In a typical floating point number system a non-zero number x is stored in the form

$$x = \sigma \cdot (.a_1a_2a_3\cdots a_t)_\beta \cdot \beta^e,$$

   where $\sigma = +1$ or -1, $a_1 \neq 0$, $0 \leq a_i \leq \beta - 1$, $t = 53$, $\beta = 2$, and $-1023 \leq e \leq 1024$.

   (a) Find the greatst and smallest positive numbers and the unit roundoff.

   (b) Which of the following are the numbers in this typical floating point number system?

$$10, \quad 1 + 2^{-53}, \quad 1 - 2^{-53}, \quad 2^{1024}.$$