

# CSE 330 [Numerical Methods]

## Assignment 01 [Spring 2025]

---

### Instructions for submission: [Handwritten submission]

- Write your Name, Student\_ID, Section No. in the cover page of the assignment.
  - Mark the answers properly for each corresponding question.
  - There is no late submission.
- 

1. In the classes, we discussed three forms of floating number representations, as shown below,

**Standard/General Form:**  $F = \pm(0.d_1d_2d_3 \cdots d_m)\beta^e$ ,

**Normalized Form:**  $F = \pm(1.d_1d_2d_3 \cdots d_m)\beta^e$ ,

**Denormalized Form:**  $F = \pm(0.1d_1d_2d_3 \cdots d_m)\beta^e$

Now, let's take,  $\beta = 2$ ,  $m = 4$  and  $-2 \leq e \leq 6$ . Based on these, answer the following:

- (a) (6 marks) What are the **maximum/largest and minimum/smallest** numbers that can be stored in the system by these three forms defined above (express your answer in decimal values)?
- (b) (3 marks) What are the **non-negative minimum/smallest** numbers that can be stored in the system by the three forms defined above (express your answer in decimal values)?
- (c) (3 marks) How many **numbers in total/ possible combinations** can be represented by this system? Find this separately for each of the three forms above including the **negative numbers**.

2. Consider the **real number**  $x = (5.625)_{10}$

- (a) (2 marks) First convert the decimal number  $x$  in binary format.
- (b) (3 marks) Find **fl(x)** if you store it in a system with  $m = 3$  using the **denormalized** form of floating point representation.
- (c) (6 marks) Now convert back to decimal form the stored value you obtained in the previous part, and calculate the **Rounding Error** and **Machine Epsilon**.

3. Consider the quadratic equation,  $x^2 - 60x + 1 = 0$  Below, calculate **up to 6 significant** figures.

- (a) (4 marks) Find out where the loss of significance occurs when you calculate the roots?
- (b) (3 marks) Evaluate the correct roots such that loss of significance does not occur.