Lecture 1

Number Representation on a finite word-size machine

Numbers

- Integer
- Real
- Rational
- Irrational
- Real Number Representation:
  - Fixed Point
  - Floating Point

Rational Numbers

- A rational number = Integer 1/ Integer 2
- Example:

3/5 = 0.600000000000000

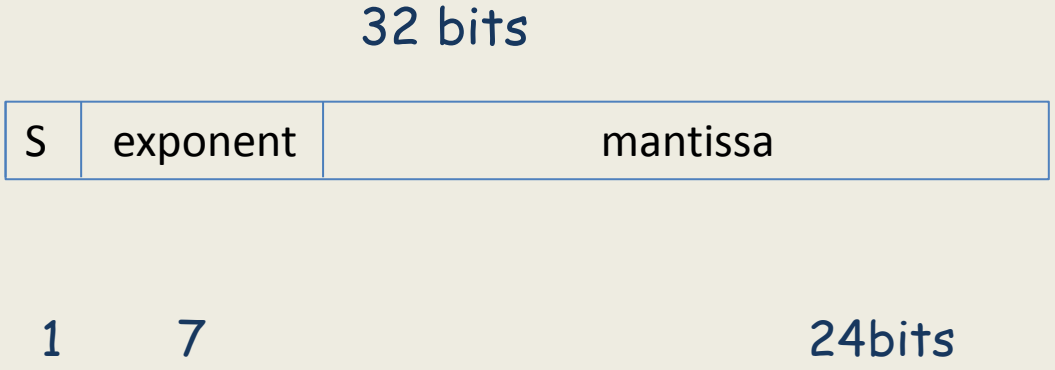
Can be represented accurately on a fixed size register or memory word.

However, 7/3 = 2.33333…333 has to be approximated.

Therefore, some rational numbers can be accurately represented on a computer memory while others have to be approximated.

Real Numbers

- Only a relatively small subset of real numbers can be accurately represented on a fixed size memory word. Some approximation is always required.
- Example: A 32-bit floating point data representation



Example on Floating Point Number Representation

- $(-1)^{-S} \cdot (1 + f) \cdot \text{base}^{e-a}$
- Example : base = 16, bias a = 64 and number is as follows:

Decimal number is calculated as follows:

0	1000010	1011 0011 0000 0100 0000 0000

$[1+ (2)^{-1} + (2)^{-3} + (2)^{-4} + (2)^{-7} + (2)^{-8} + (2)^{-14}] .16^{66-64}$

$= 1.699279785 . 16^{2} = 435.0156001$ (with 1 added)

or $= 0.699279785 . 16^{2} = 179.0115625001$ (without 1)

What is the next larger and the previous smaller numbers that can be represented on this machine?