

# **ARM ASSIGNMENT 3**

**P SAI PHANINDRA**

**MT2017515**

## **1. PRECISION AND EXPONENT**

Precision is how close the number is to the actual value.

Sign bit represents positive or negative number

Exponent means the no of places the decimal have to move.

there must be three values required to make a value precise. sign and exponent remain accurate leaving the precision dependent on fractional part.

fractional part consists of binary part after the leading but the size of fractional part is only 22 bits and the value is less precise.

EX : 0000 1000 1101 1101 1101 1101 1110 0111 is 32 bit number

Remaining bits are removed and the value is less precise because of rounding.

Exponent part is same for precise as well as precise number which is 27.

## **2. NORMAL/SUBNORMAL VALUES**

exponent can vary from -126 to 127 with the significand part.

for example consider 1101

to get leading 1 our exponent must be less than -127 which is not valid.

we can allow leading 0 instead and the number using exponent as -127. these values with leading 0 are subnormal values.

### **3. ROUNDING OFF**

- a. rounding off to its nearest even integer.

ex : 17.8 to 18

- b. rounding off towards other side of zero

ex: 24.5 to 25

- c. Rounding off towards zero

ex: 56.9 to 56

- d. Rounding off towards minus infinity. it is the same as rounding off towards near zero

ex : 14.3 to 14

- e. Rounding towards plus infinity : away from zero.

ex : 14.8 to 15