Question 1.

In Floating point representation, we have three components

1.The Sign Bit
2.Exponent
3.Fractional Part

-

Precession is one the prime attribute of any Floating-Point Representation,

1.Does any of the above three components play a role in the defining the Precession of the number? If so which are the component or Components which play the role in defining precession and how ? Explain this with example in your own words

Floating Point Representation is used for representing real number with some approximation. It has a trade of between precision and range because every number cannot be represented in the form (significand)* base^exponent. It is used to represent very large real numbers. (1.) Precision is one of the prime attribute. The fractional part plays main role in defining the precision of the number. The length of the fractional part (or significand) determines the precision of the number. Example: Consider base 10 12021.414 --> This number is represented as 1.2021414 * 10^4 here the significand is 8bits long If we consider the significand to be 6bits(say) then the number would be 1.20214 * 10^4. The former has a better precision than the later. Hence the fractional part effects the precision of a number.

2.What is Normal and Subnormal Values as per IEEE 754 standards explain this with the help of number line

Consider a 32bit representation, we have 1 sign bit, 8bits in exponent part and 23bits in fractional part of the representation. A value is said to be normalized when the number is represented with having a single bit before the decimal. e.g decimal number 123.5 is represented in binary as 1111011.1 When the number is normalised it is represented as 1.1110111 * 10^6 Subnormal numbers are under the denormal number category. A number which has a magnitude smaller than the smallest normal number is called a subnormal number.

3.IEEE 754vv defines standards for rounding floating points numbers to a represent able value. There are five methods defines by IEEE for this – Take time and understand what these five methods and explain it in your words using diagrams, illustrations of your own.

IEEE defines five methods for rounding of floating point numbers.

1. Round to nearest, ties to even – rounds to the nearest value: rounded to nearest value that has 0 as least significant bit
2. Round to nearest, ties away from zero: rounded to nearest value above given number for positive number and below for negative numbers
3. Round toward 0: rounding is directed towards zero
4. Round toward +∞: rounding is directed towards +ve infinity
5. Round toward -∞: rounding is directed towards -ve infinity

```
   e.g   +14.4     -14.4     +15.4
```
Round to nearest, ties to even : +14.0 -14.0 +16.0

Round to nearest, ties away from zero : +15.0 -15.0 +16.0

toward 0 : +14.0 -14.0 +15.0

toward +∞ : +15.0 -14.0 +16.0

toward -∞ : +14.0 -15.0 +15.0