Experiment No-6

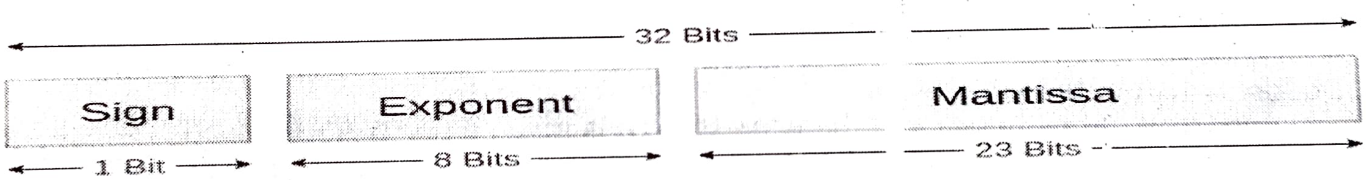
~~Q 3A. For 132.65~~ obtain the IEEE 754 standards of Single precision and Double Precision format for $(85.125)_{10}$?

The IEEE Standard for Floating-Point Arithmetic (IEEE 754) is a technical standard for floating-point computation which was established in 1985 by the **Institute of Electrical and**

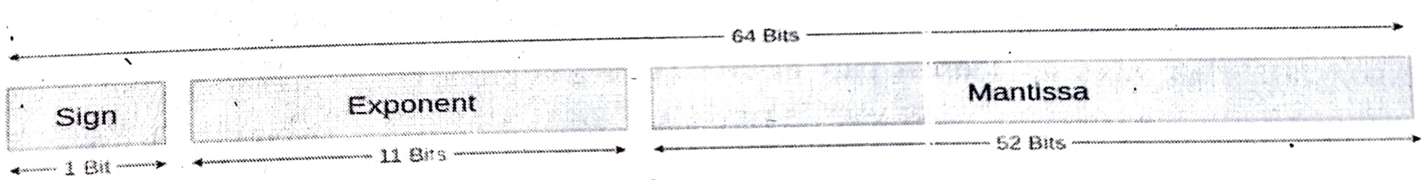
Electronics Engineers (IEEE). The standard addressed many problems found in the diverse floating point implementations that made them difficult to use reliably and reduced their portability. IEEE Standard 754 floating point is the most common representation today for real numbers on computers, including Intel-based PC's, Macs, and most Unix platforms.
There are several ways to represent floating point number but IEEE 754 is the most efficient in most cases. IEEE 754 has 3 basic components:

1. **The Sign of Mantissa –**
   This is as simple as the name. 0 represents a positive number while 1 represents a negative number.
2. **The Biased exponent –**
   The exponent field needs to represent both positive and negative exponents. A bias is added to the actual exponent in order to get the stored exponent.
3. **The Normalised Mantissa –**
   The mantissa is part of a number in scientific notation or a floating-point number, consisting of its significant digits. Here we have only 2 digits, i.e. O and 1. So a normalised mantissa is one with only one 1 to the left of the decimal.

IEEE 754 numbers are divided into two based on the above three components: single precision and double precision.



Single Precision
IEEE 754 Floating-Point Standard



Double Precision
IEEE 754 Floating-Point Standard

TYPES	SIGN	BIASED EXPONENT	NORMALISED MANTISA	BIAS
Single precision	1(31st bit)	8(30-23)	23(22-0)	127
Double precision	1(63rd bit)	11(62-52)	52(51-0)	1023

Example –
85.125

85 = 1010101

0.125 = 001

85.125 = 1010101.001

=1.010101001 x 2^6

sign = 0

1. Single precision:

biased exponent 127+6=133

133 = 10000101

Normalised mantisa = 010101001

we will add 0's to complete the 23 bits

The IEEE 754 Single precision is:

$= 0\ \underbrace{10000101}_{\text{Exp}}\ \underbrace{01010100100000000000000}_{\text{Mantisa}}$

This can be written in hexadecimal form **42AA4000**

2. Double precision:
biased exponent 1023+6=1029
1029 = 10000000101
Normalised mantisa = 010101001
we will add 0's to complete the 52 bits

The IEEE 754 Double precision is:
$= 0\ \underbrace{10000000101}\ \underbrace{0101010010000000000000000000000000000000000000000000}$
This can be written in hexadecimal form **4055480000000000** Mantissa

Click reset to clear and re-enter the values

Simulations

DECIMAL NUMBER	85.125
BITS FOR EXPONENT	8

Submit Reset

RESULTS

8-bit binary	0 1 0 0 0 0 1 0 1
Binary Representation Of Integeral Part	1010101
Binary Representation Of Fractional Part	001
Binary Representation of the Number	1010101.001
Normalised Representation of the Number	1. X 2 power6

Sign	0	Bias	127
Mantiss		Expone	133

Click reset to clear and re-enter the values

Simulations

DECIMAL NUMBER	85.125
BITS FOR EXPONENT	11

Submit Reset

RESULTS

8-bit binary	0 1 0 0 0 0 0 0 0 1 0 1
Binary Representation Of Integeral Part	1010101
Binary Representation Of Fractional Part	001
Binary Representation of the Number	1010101.001
Normalised Representation of the Number	1. X 2 power6

Sign	0	Bias	1023
Mantiss		Expone	1029