Welcome to ACE Engineering Academy - online live class

Subject: **Computer Organization and Architecture**

Faculty: **Y.V. Ramaiah**

Subject

Computer organization & Architecture

Chapters (Topics)

I. Computer Arithmetic

II. Memory Organization

III. Secondary Memories

IV. Basic processor organization and Design

V. Pipeline organization

VI. Control unit Design

VII. IO Organization

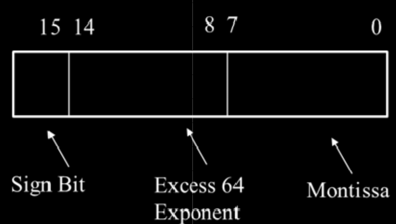
Chapter - 1
Computer Arithmetic

1. Signed binary data representation (Integer)
2. Overflow concept in signed 2's complement representation
3. Different Arithmetical operations
4. Booth's Algorithm
5. Floating point representation
6. IEEE standards for floating point representation
7. Ripple carry Adder
8. Carry lookahead generator and Adder

Q. The data is given below. Solve the problems and choose the correct answer.

0.239×2^{13}

Binary



The normalized representation for the above format is specified as follows. The mantissa has an implicit preceding the binary (radix) point. Assume that only 0's are padded in while shifting a field. The normalized representation of the above (0.239×2^{13}) is

(a) 0A 20 (b) 11 34 (c) 4D D0 (d) 4A E8

$0{\cdot}00111101 \times 2^{13}$ Non Normalized

$0{\cdot}11110100 \times 2^{11}$ Explicit

$1{\cdot}\underline{1110100 0} \times 2^{10}$ Implicit

$M = 11101000$

$e = 10, \; b = 64$

$E = 74 \quad S = 0$

1001010

S	E	M		
0 100	1010	1110	1000	
4	A	E	8	

(4AE8)

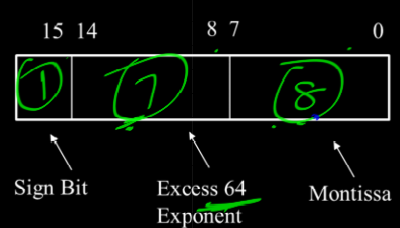
Q. The data is given below. Solve the problems and choose the correct answer.

$b = 64 = 2^6 = 2^{7-1}$

$K = 7$

$0{\cdot}239 \times 2^{13}$

Binary Mantissa



Mantissa is a pure fraction in signed magnitude form. The decimal number 0.239×2^{13} has the following hexadecimal representation without normalization and rounding off

(a) 0D 24 (b) 0D 4D

(c) 4D 0D (d) 4D 3D

0

0.239×2^{13} $\quad S = 0$

$= 0.\overset{M}{00111101} \times 2^{13}$

$e = 13,\ b = 64,\ E = 77$
$= 1001101$

S	E	M
0	1001101	00111101

0100	1101	0011	1101
4	D	3	D

$.239 \times 2$
0.478×2
0.956×2
1.912×2
1.824×2
1.648×2
1.296×2
0.592×2
1.184

Q. The value of a float type variable is represented using the single-precision 32-bit floating point format of IEEE-754 standard that uses 1 bit for sign, 8 bits for biased exponent and 23 bits for mantissa. A float type variable X is assigned the decimal value of -14.25. The representation of X in hexadecimal notation is

(a) C1640000H ✓ (b) 416C0000H

(c) 41640000H (d) C16C0000H

2014 GATE

2022 GATE

MCQ

Q. Consider three floating point numbers A, B and C stored in registers RA, RB and RC, respectively as per IEEE-754 single precision floating point format. The 32-bit content stored in these registers (in hexadecimal form) are as follows.

-12 +36 +12

R_A= 0xC1400000	R_B = 0x42100000	R_C = 0x41400000

Which one of the following is FALSE?

(A) A + C = 0 TRUE

(B) C = A + B False

(C) B = 3C TRUE

(D) (B − C) > 0 TRUE

R_A = 1 | 1000 0001 | 0100 000...0 (S | E = 130 | M), b = 127

Value = $(-1)^1 * 1.100...0 \times 2^3$

= (−) 1100.00 × 2^0 = −12

R_B = 42100000

= 0 | 1000 0100 | 001 000...0 (S | E = 132, b = 127 | M)

Value = $(-1)^0 \times 1.00100 \times 2^5$

= (+) 100100.0 × 2^0

= +36

R_C = 41400000

= 0 | 1000 0010 | 100000...0 (S | E = 130 | M)

Value = $(-1)^0 * 1.1000 \times 2^3$ = + 1100 × 2^0

= +12

Ex) A 16 bit Register is used to store floating point data with Excess 32 technique.

S	E	M
1	6	9

16

$b = 32 = 2^5 = 2^{K-1}$

$K = 6$

1Q) No. of bits needed for 'M' field is ~~9~~

2Q) No. of bits needed for 'E' field is ~~6~~

3Q) store $(17.5)_{10}$ in the above Register and Express the value in Hexadecimal.

$S = 0$ $\qquad 17.5 = 10001.1 \times 2^0$

$17.5 = 10001.1 \times 2^0$

$S = 0$ $\qquad = 0.100011 \times 2^5, \; e = 5, \; b = 32$

$= 0.\underline{100011000} \times 2^5 \qquad E = 37 = 100101$

$M = 100011000$

$\longleftarrow$ 16 bit $\longrightarrow$

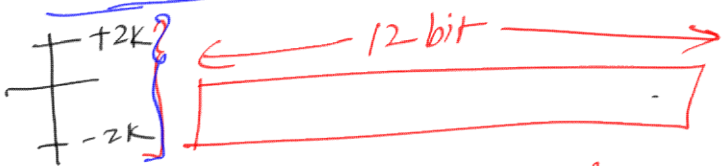
S	E	M
0	100101	100011000

0100	1011	0001	1000
4	B	1	8

$= (4B18)_{16}$

Explanation, How Floating point Representation provides more Range than Integer Representation

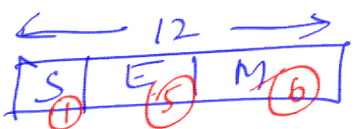
$2^{10} = 1024 \cong 1K$



Signed Integer ; Range is -2^{11} to $+2^{11}-1$

$-2K$ to $+2K$

$-2K$ to $+2K$
total, 4K



$b = 16 = 2^{K-1} = 2^4$

$K = 5$

Max. Value to be Stored

$S = \pm = 0/1$

$E_{max} = 11111 = 31$

$= \pm\, 0.111111 \times 2^{31-16}$

$= \pm (1 - 2^{-6}) \times 2^{15}$

$\cong \pm 1 \times 2^{15}$ $\cong \pm 2^5 \times 2^{10}$

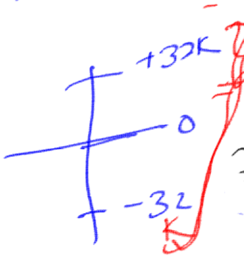
$\cong \pm 32K$

$e = E - b = 15$

$E = e + b$

$-32K$ to $+32K$

$= 64K$



$E = e + b$

$31 = e + 16$, $e = 15$

Normalization:– It is the process of maintaining the non zero digit in the MSB of the Mantissa.

MSB

Normalized 'M' = $0.\boxed{1}xxxx\ldots x$

Non Normalized 'M' = $0.\boxed{0}xxx\ldots x$

Normalization is used to reduce the size of the 'M' and it provides accurate result

Ex

S	E	M
1	6	5

$\longleftarrow 12 \longrightarrow$

$b = 32 = 2^5 = 2^{k-1}$

$k = 6$

Result

Non-Normalized

$0.\underline{00011101} \times 2^6$

$M = 00011101$

$e = 6$ (8 bit)

Normalized

$= 0.11101 \times 2^3$

$M = 11101$ 5 bit

No Accuracy problem

Normalization Types

→ There are 2 types

(i) Explicit

Normalized bit is positioned in the MSB of the Mantissa (In default).

(ii) Implicit

The normalized bit is Hidden i.e. the normalized bit is shifted to Integer position (IEEE standards follow this)

Ex

0.0001011×2^5

MSB

0.1011000×2^2 — Explicit Normalized 'M'

$M = 1011000$, there is no Hidden bit

1.0110000×2^1 — Implicit Normalized M

Hidden bit

$M = 0110000$

Implicit Normalization provides more Range than Explicit Normalization.

Explicit Normalized : $0.1xxxxx \times 2^{e}$

Implicit Normalized $1.xxxxx \times 2^{e}$

Hidden

Expression Values

Explicit Normalization

$(-1)^S * 0.M \times 2^{E-b}$ ✓

Implicit Normalization

$(-1)^S * 1.M \times 2^{E-b}$ ✓

IEEE Standards use this

Implicit Normalization provides more Range than Explicit Normalization Value for the given Register size

---

$b = 16 = 2^{K-1}$

$K = 5$

12

S	E	M
1	5	6

$E_{max} = 11111_2 = 31$

Explicit Normalized Value

Max Value $\pm 0.111111 \times 2^{31-16}$

$\simeq \pm 1 \times 2^{15}$

$\simeq \pm 32K$

Implicit Normalized Value

$(-1)^S * 1.M \times 2^{E-b}$

Max Value $= \pm 1.111111 \times 2^{15}$

$\frac{1}{2} \frac{1}{4} \frac{1}{8} \frac{1}{16} \frac{1}{32} \frac{1}{64}$

$= \pm 1.999 \times 2^{15}$

$\simeq \pm 2 \times 2^{15} \simeq \pm 2^{16}$

$\simeq \pm 64K$

Q) The mantissa size is 6 bit, and it is in signed magnitude form with Explicit Normalization. The decimal Value Range is

a) 0 to $(1-2^{-6})$ b) 0.5 to $(1-2^{-6})$

c) 1 to $(1-2^{-6})$ d) (0 to 1)

$M_{min} = 0.\underset{\frac{1}{2}}{1}00000$ (6 bit)

$(0.5 \text{ to } 1-2^{-6})$

$M_{max} = 0.111111 = (1-2^{-6}) \cong 1$

Ex) A 12 bit Register is used to store floating point Representation for storing $(-13.25)_{10}$ with biasing value 16

1Q) The value in the Register when it uses Explicit Normalization is ___

2Q) The value in the Register when it uses implicit Normalization is ___

$b = 16 = 2^4 = 2^{5-1}$

$K = 5$

S (1)	E (5)	M (6)

-13.25

$S = 1$

Explicit Normalization

13.25

1101.01×2^0

0.110101×2^4

$M = 110101$

$e = 4, \ b = 16$

$E = 20$

10100

$S = 1$

S	E	M
1	10100	110101

1101	0011	0101
D	3	5

$(D35)_{16}$

0.110101×2^{4}

Implicit Normalized form

1.101010×2^{3}

$S = 1$

$M = 101010$

$e = 3,$

$E = 3 + 16 = 19$

10011

S	E	M
1	10011	101010

1100 | 1110 | 1010

C | E | A

(CEA)

Limitations for Normalized general Representation.

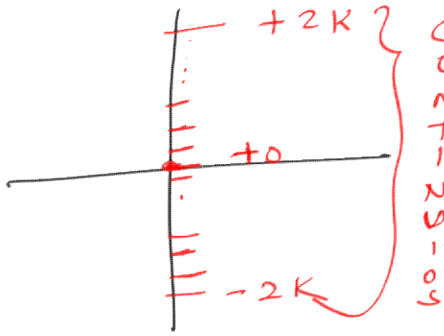
→ Since zero can't be normalized, it is not possible to Represent '0' in General Notation.

→ The Range is discontinuous.

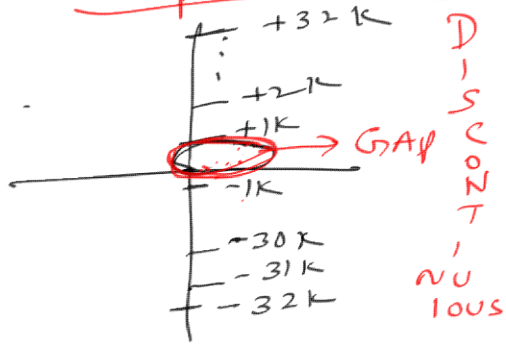
Hence IEEE standards are adopted

Signed Integer Range

12 bit



General Normalized Floating point Representation



IEEE standards

→ They follow only Implicit Normalization

→ let Exponent field size is K bits,

$b = 2^{k-1} - 1$

→ It is possible to Represent ± 0, $\pm$ infinitive and Not a number (NAN)

→ Emin and Emax Values are Reserved for Representing these Special Values.

$\frac{0}{x} = 0$; $\frac{x}{0} = \infty$

and $\frac{0}{0}$, $\sqrt{-1}$ = NAN

$E_{min} = 0$, E_{max} Values are Reserved for Representing these Special Values

IEEE Representations

Single precision
32 bit size
$b = 127$

Double precision
64 bit size
$b = 1023$

Single precision (32 bit)

$\longleftarrow 32 \longrightarrow$

S	E	M
1	8	23

$E = e + b$, $b = 127$

$E_{min} = 00000000_2 = 0$

$b = 127 = 2^{K-1} - 1 = 127$

$E_{max} = \underset{128}{1}\,\underset{64}{1}\,\underset{32}{1}\,\underset{16}{1}\,\underset{8}{1}\,\underset{4}{1}\,\underset{2}{1}\,\underset{1}{1}{}_2 = 255$

$2^{K-1} = 128$

$2^{K-1} = 2^7$

$K-1 = 7$

$K = 8$

Emin and Emax are Reserved for Special Values

IEEE 32 bit Notation

$$\text{Expression Value} = (-1)^S * 1.M \times 2^{E-127}$$

$$= (-1)^S * 1.M * 2^{E-b}$$

(1) S	E (8)	M (23)	Representation
0 / 1	00000000 E=0	000.......0 M=0	± 0
0 / 1	1111 1111 255	000.....0	± ∞ (infinitive)
0 / 1	1111 1111 (255)	M ≠ 0	NAN
0 / 1	1 ≤ E ≤ 254	x x x. ...x	Implicit Normalized Number (Gener Number)
0 / 1	0000000 E=0	M ≠ 0	Denormal Number ≅ 0

VVImp

IEEE Standards are also known as IEEE 754 standards.

Q) 0×80000000 is used to Represent IEEE Single precision floating data; it's value is

a | Normalized +0 ✗
b | Normalized −0 ✗
c | Special Value +0 ✗
d | Special Value −0 ✓

S = 1

(−0)

1 | 0000 0000 0 | 0000000 ... 0

S | E = 0 | M = 0.

Q) Represent -14.25_{10} in IEEE single precision format and Express the value in Hexadecimal

S = 1 : 14.25

$= 1110.01 \times 2^0$

$= 1.11001 \times 2^3$

Hidden

M = 1100100000.

$e = 3, b = 127$

$E = 130 = 10000010$

S	E	M
1	10000010	1100100 ... 0

1100 | 0001 | 0110 | 0100 | 0000

C 1 6 4 0000

$(C1640000)_{16}$

ISRO IT GATE :-

Q) $C1D00000_{16}$ is represented in IEEE Single precison format, it's decimal Value is ____

$$\overset{S}{1}\,|\,\overset{128}{1}000001\overset{2\ 1}{1}\,|\,1010000\cdots0$$

$E = 131$

$b = 127$

$e = 4$

$M = 1010000\cdots0$

$$\text{Value} = (-1)^S * 1.M \times 2^{E-b}$$

$$(-)\ 1.1010.00 \times 2^4 = (-)\overset{16\,8\,4\,2\,1}{11010}.00_2 \times 2^0$$

$(-)\ 26$