CH-10

Topics to be covered from Chapter-7

Lecture -27: Figure 7.2, 7.3

Lecture-29 & 30:Table 7.1, Figure 7.4, 7.8, 7.9,7.10

Lecture-31 (OPTIONAL): Figure 7.12, 7.14, 7.15

Topics to be covered from Chapter-10

1. Disadvantages of multiplication of unsigned binary integers
2. Twos complement Multiplication (Booth's algorithm) (Figure 10.12, Figure 10.13
3. Multiplication of following numbers using 4-bit binary integers and Booth's algorithm
   a. 7x3
   b. 7x(-3)
   c. (-7)x(3)
4. Floating point Representation (Figure 10.18 only)
5. IEEE standard for binary floating-point Representation (Figure 10.21 (a), Figure 10.21(b) only)
6. Express the following numbers in IEEE 32-bit floating-point format(page number 370)
   a. -1.5
   b. 0.384
   c. -1/32 = −0.031
7. The following numbers use the IEEE 32-bit floating-point format. What is the equivalent decimal value?(page number 370)
   a. 1 10000011 11000000000000000000000
   b. 0 01111110 10100000000000000000000

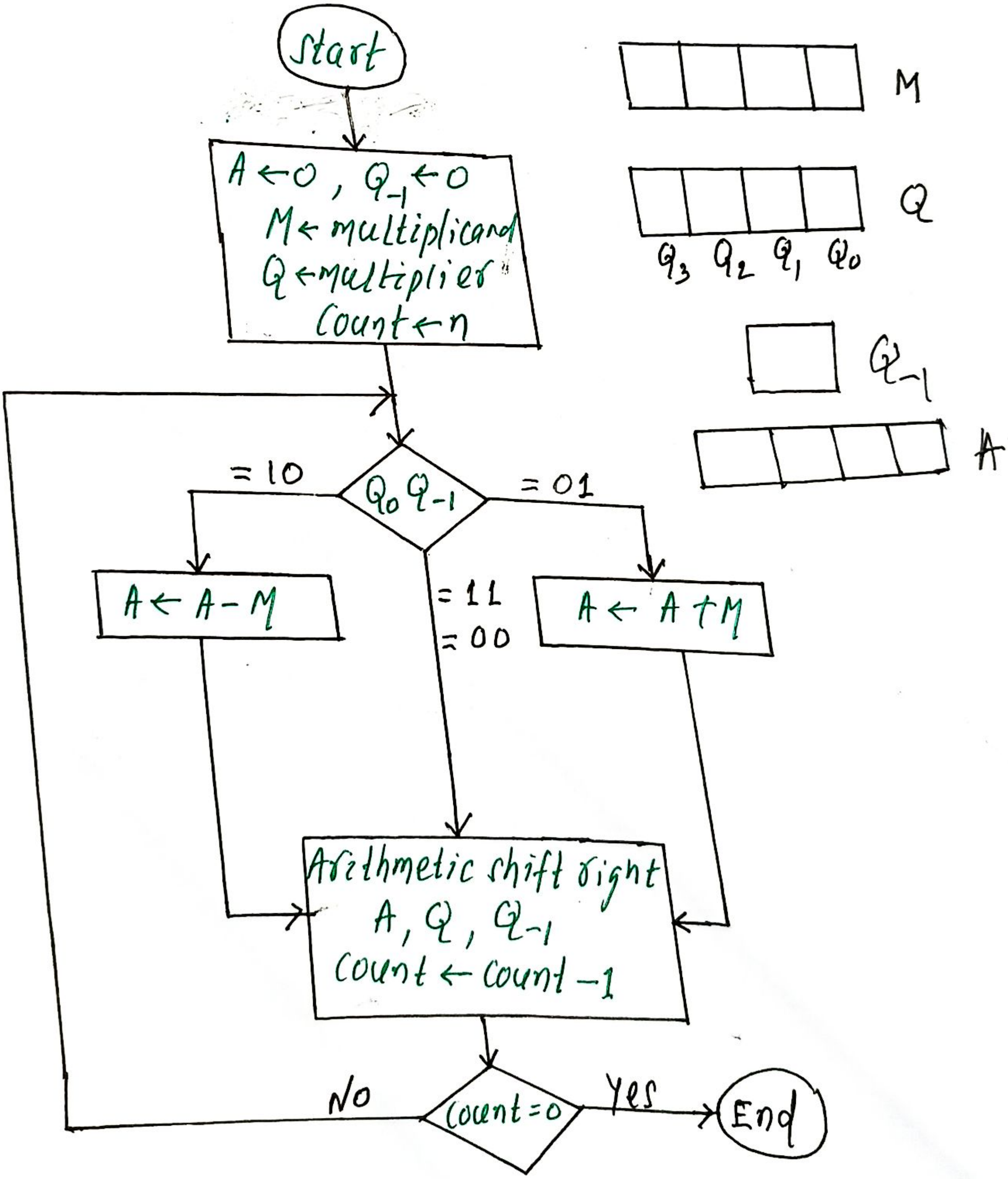
10.3 Booth's Algorithm for twos complement multiplication

M = multiplicand, Q = multiplier

A is 4 bit register, Q_{-1} is a single bit register.

The result is stored in A & Q registers.

A & Q_{-1} are intialized at 0, & n is no bits



Q.1 Multiplication of (7X3) using 4-bit binary integers based on the Booth's algorithm.

Solution

$M = 7 = 0111$, $-M = 1000 + 1 = 1001$

$Q = 3 = 0011$

A ← A−M
0000
+1001
1001

A ← A+M
1110
+0111
10101

A	Q	Q_{-1}	M		
0000	0011	0	0111	Initial Value	
1001	0011	0	0111	A ← A−M	1st cycle
1100	1001	1	0111	shift	
1110	0100	1	0111	shift	2nd cycle
0101	0100	1	0111	A ← A+M	3rd cycle
0010	1010	0	0111	shift	
0001	0101	0	0111	shift	4th cycle

The result is $AQ = (0001\,0101)_2 = (21)_{10}$

Q.2 Multiply 7×(−3), using Booth's algorithm

$M = 7 = 0111$, $-M = 1000 + 1 = 1001$

$Q = -3 = 1100 + 1 = 1101$

A ← A−M
+0000
1001
1001

A ← A+M
1100
+0111
10011

A ← A−M
0001
1001
1010

A	Q	Q_{-1}	M		
0000	1101	0	0111	Initial value	
1001	1101	0	0111	A ← A−M	1st cycle
1100	1110	1	0111	shift	
0011	1110	1	0111	A ← A+M	2nd cycle
0001	1111	0	0111	shift	
1010	1111	0	0111	A ← A−M	3rd cycle
1101	0111	1	0111	shift	
1110	1011	1	0111	shift	4th cycle

The result is $AQ = 11101011 = -21$

Q.3 Multiplication of (-7 X 3), using 4-bit integers based on the Booth's algorithm;

Solution →

$M = -7 = \begin{array}{r} 1000 \\ +\ 1 \\ \hline 1001 \end{array}$ $Q = 3 = 0011$

$\Rightarrow -M = 7 = 0111$

$A \leftarrow A - M$

$\begin{array}{r} 0000 \\ +0111 \\ \hline 0111 \end{array}$

A	Q	Q_{-1}	M		
0000	0011	0	1001	Initial Value	
0111	0011	0	1001	$A \leftarrow A - M$	1st cycle
0011	1001	1	1001	shift	
0001	1100	1	1001	shift	2nd cycle
1010	1100	1	1001	$A \leftarrow A + M$	3rd cycle
1101	0110	0	1001	shift	
1110	1011	0	1001	shift	4th cycle.

$A \leftarrow A + M$

$\begin{array}{r} 0001 \\ +1001 \\ \hline 1010 \end{array}$

The result of $AQ = (1110\ 1011)_2 = (-21)$

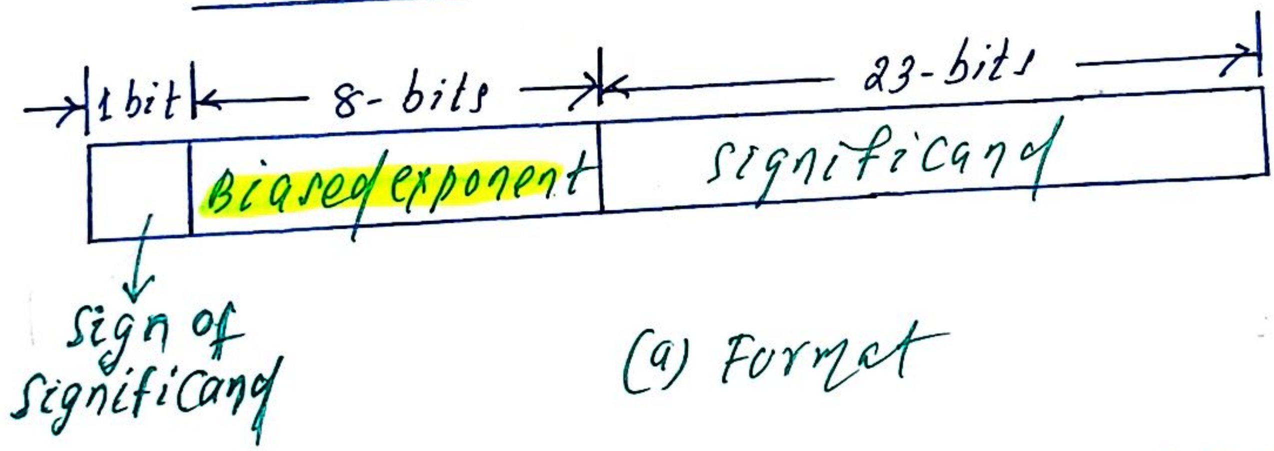
10.4 Floating point representation →

Fig 10.18 (a) shows a 32-bit floating point format.

→ The leftmost bit is sign bit, 0 = positive, 1 = negative.

→ The exponent value is stored in next 8-bit. The bias of this 8-bit field is $(2^{k-1}-1) = (2^{8-1}-1) = (2^7-1) = 127$. where k is no of bits in binary exponent.

→ The final portion of this format is significand. it is 23 bits.



(a) Format

0 1011 0001 011110000110000000011111

1 0000 1001 1111 1111 0000 0000 111000

(b) example.

Fig 10.18 Typical 32-bit Floating point Format.

IEEE standard for binary floating-point

→ The IEEE 754 format have 3-basic binary format of length 32, 64 & 128 bits with exponents of 8, 11, & 15 bits, respectively.

1 bit	8-bits	23-bits
sign bit	Biased exponent	Significand

(a) 32-format

1 bit	11-bits	52-bits
sign bit	Biased exponent	significand

(b) 64-format

1 bit	15-bits	112-bits
sign bit	bised exponent	significand.

(c) 128-format

Fig 10.21 IEEE 754 Format

Q.6 Find IEEE 32 bit floating point of -1.5

Soln Start with positive version of -1.5:

$|-1.5| = 1.5$

1st find the binary no of $(1.5)_{10}$?

For $(1)_{10} = (1)_2$

$$\begin{array}{r|l} 2 & 1 \rightarrow 1 \uparrow \\ 2 & 0 \rightarrow 0 \\ & 0 \end{array}$$

Then $(0.5)_{10} = (0.10)_2$

$.5 \times 2 \rightarrow 1.0 \rightarrow 1 \downarrow$
$.0 \times 2 \rightarrow 0.0 \rightarrow 0$

Now $(1.5)_{10} = (1.1000\ 0000\ 0000\ 0000\ 000)_2$

Normalization of binary number:

→ shift the decimal point to the right or left, so that only one non zero digit remains to the left of it.

→ Hence Normalized binary representation of

$1.1000\ 0000\ 0000\ 0000\ 0000 \times 2^0$

Hence, sign: 1 (a negative number)

Exponent (unadjusted): 0

Significand (not normalized): $1.1000\ 0000\ 0000\ 0000\ 0000$

Exponent (adjusted) $= 0 + (2^{k-1}-1) = 0 + (2^{7}-1) = (0+127)_{10}$
$= (127)_{10}$

Now $(127)_{10} = (0111\ 1111)_2$

2	127	→ 1
2	63	→ 1
2	31	→ 1
2	15	→ 1
2	7	→ 1
2	3	→ 1
2	1	→ 1
	0	

Normalized significant

Significand (not normalized) = 1.1000 0000 0000 0000 0000
Significand (normalized) = 100 0000 0000 0000 0000 0000

Conclusion:
Sign (1 bit) = 1 (a negative no)
Exponent (8 bits) = 0111 1111
Significand (23 bit) = 100 0000 0000 0000 0000 0000

∴ IEEE 32 bit floating point is
1 - 0111 1111 - 100 0000 0000 0000 0000 0000

Q.7 Find the IEEE 32 bit floating point of 0.384

Solution To find the binary no of $(0.384)_{10}$

$(0)_{10} = (0)_2$

$(.384)_{10} = (.0110\ 0010\ 0100\ 1101\ 1101\ 00101)_2$

$.384 \times 2 = 0.768 \rightarrow 0$

$.768 \times 2 = 1.536 \rightarrow 1$

$.536 \times 2 = 1.072 \rightarrow 1$

$.072 \times 2 = 0.144 \rightarrow 0$

$.144 \times 2 = 0.288 \rightarrow 0$

$.288 \times 2 = 0.576 \rightarrow 0$

$.576 \times 2 = 1.152 \rightarrow 1$

$.152 \times 2 = 0.304 \rightarrow 0$

$.304 \times 2 = 0.608 \rightarrow 0$

$.608 \times 2 = 1.216 \rightarrow 1$

$.216 \times 2 = 0.432 \rightarrow 0$

$.432 \times 2 = 0.864 \rightarrow 0$

$.864 \times 2 = 1.728 \rightarrow 1$

$.728 \times 2 = 1.456 \rightarrow 1$

$.456 \times 2 = 0.912 \rightarrow 0$

$.912 \times 2 = 1.824 \rightarrow 1$

$.824 \times 2 = 0.648 \rightarrow 0$

$.648 \times 2 = 1.296 \rightarrow 1$

$.296 \times 2 = 0.592 \rightarrow 0$

$.592 \times 2 = 1.184 \rightarrow 1$

$.184 \times 2 = 0.368 \rightarrow 0$

$.368 \times 2 = 0.736 \rightarrow 0$

$.736 \times 2 = 1.472 \rightarrow 1$

$.472 \times 2 = 0.944 \rightarrow 0$

$.944 \times 2 = 1.888 \rightarrow 1$

$\therefore (0.384)_{10} = (0.0110\ 0010\ 0100\ 1101\ 1101\ 0010\ 1)_2$

Normalization of binary no:

→ shift the decimal point to the right or left, so that only one non-zero digit remains to the left.

→ Normalized binary no:

$1.1000\ 1001\ 0011\ 0111\ 0100\ 101 \times 2^{-2}$

Hence

Sign : 0 (positive no)

Exponent (unadjusted) = −2

Significand (not Normalized) = 1.1000 1001 0011 0111 0100 101

Exponent (adjusted) = $-2 + 127 = (125)_{10}$

Now $(125)_{10} = (0111\ 1101)_2$

Normalized significand:

Significand (not normalised) =

~~1.~~1000 1001 0011 0111 0100 101

Significand (normalized) =

1000 1001 0011 0111 0100 101

Conclusion:

Sign = 0

Exponent = 0111 1101

Significand = 1000 1001 0011 0111 0100 101

∴ IEEE 32 bit floating point is

0-0111 1101-1000 1001 0011 0111 0100 101

Q.8 Find the IEEE 32 bit floating point of $-1/32$

Solution $-\frac{1}{32} = -0.031$

Q.8 IEEE 32 bit floating point number
1 10000011 11000000000000000000000,
Find the equiva nt decimal value?

Solution →

→ The 1st bit indecates sign, 1= negative, 0= positive.

Sign bit = 1

→ The next 8-bits contain the exponent.

exponent = 1000 0011

→ last 23 bit contain the significand.

Significand = 11000000000000000000000

→ The exponent $(1000\ 0011)_2 = 1\times2^7 + 1\times2^1 + 1\times2^0$
$= 128+2+1 = (131)_{10}$

→ Adjusted exponent:

$131 - 127 = 4$

→ Significand:

$(110\ 0000\ 0000\ 0000\ 0000\ 0000)_2 = 1\times2^{-1} + 1\times2^{-2} + 0\times2^{-2} + \cdots = 0.5 + 0.25$
$= (0.75)_{10}$

The floating point decimal value:

$$(-1)^{sign} \times (1 + \text{significand}) \times 2^{\text{adjusted exponent}}$$

$$= (-1)^1 \times (1 + 0.75) \times 2^4$$

$$= -1.75 \times 2^4 = -28.0$$

$\therefore$ The equivalent decimal value is -28.0_{10}

Q·9 IEEE 32 bit floating point number
0-0111 1110-101 0000 0000 0000 0000 0000,
Find the equivalent decimal value?

Solution →

→ Sign bit = 0

→ The exponent = $(0111\ 1110)_2 = 0\times2^7 + 1\times2^6 + 1\times2^5 + 1\times2^4 + 1\times2^3 + 1\times2^2 + 1\times2^1 + 0\times2^0$

$= 64 + 32 + 16 + 8 + 4 + 2 + 0 = (126)_{10}$

→ Adjusted exponent = 126 − 127
= −1

→ significand:

$(101\ 0000\ 0000\ 0000\ 0000)_2 = 1\times2^{-1} + 1\times2^{-3} = 0.5 + 0.125$

$= (0.625)_{10}$

The floating point decimal value:

$(-1)^{sign} \times (1 + \text{significand}) \times 2^{\text{adjusted exponent}}$

$(-1)^0 \times (1 + 0.625) \times 2^{-1} = 1.625 \times 0.5$

$= (0.8125)_{10}$

∴ The equivalent decimal value is 0.8125_{10}