## Number Representation and Arithmetic Circuits

Chapter 5



#### Chapter Objectives

000

- Representation of numbers in computers
- Circuits used to perform arithmetic operations
- Performance issues in large circuit



#### Contents



- Number Representations in Digital Systems
- 2. Addition of Unsigned Numbers
- 3. Signed Numbers
- 4. Fast Adders
- 5. Multiplication
- 6. Other Number Representations
- 7. ASCII Character Code
- 8. Examples of Solved Problems



## NUMBER REPRESENTATIONS IN DIGITAL SYSTEMS



#### Number Representation



Unsigned integers

$$B = b_{n-1}b_{n-2} \cdots b_1b_0$$

$$V(B) = b_{n-1} \times 2^{n-1} + b_{n-2} \times 2^{n-2} + \dots + b_1 \times 2^1 + b_0 \times 2^0$$

$$= \sum_{i=0}^{n-1} b_i \times 2^i$$

Octal and Hexadecimal Representation

$$K = k_{n-1}k_{n-2} \cdots k_1 k_0$$

$$V(K) = \sum_{i=0}^{n-1} k_i \times r^i$$

### O Numbers in different systems O O

Decimal	Binary	Octal	Hexadecimal
00	000000	00	00
01	000001	01	01
02	000010	02	02
03	000011	03	03
04	000100	04	04
05	000101	05	05
06	000110	06	06
07	000111	07	07
08	001000	10	08
09	001001	11	09
10	$oxed{001010} 001011$	12 13	$egin{pmatrix} 0A \ 0B \end{pmatrix}$
12	001011	14	0C
13	001101	15	0D
	001110	16	0E
14 15	$ $ $00\overline{1}\overline{1}\overline{1}$	17	$0\overline{F}$
16	010000	20	10
17	010001	21	11
18	010010	22	12

Figure 5.1. Numbers in different systems.



#### Conversion from decimal to binary

Convert  $(857)_{10}$ 

```
Remainder
                           1 LSB (Least Significant Bit)
857 \div 2 = 428
428 \div 2 = 214
214 \div 2 = 107
107 \div 2 = 53
 53 \div 2 = 26
 26 \div 2 = 13
 13 \div 2 = 6
  6 \div 2 = 3
  3 \div 2 = 1
                              MSB (Most Significant Bit)
   1 \div 2 = 0
```

Result is (1101011001)<sub>2</sub>



#### OOO Octal and Hexadecimal Numbers OOO

Conversion binary number to octal and hexadecimal number

$$\left(\underbrace{\frac{10}{2}}_{2}\underbrace{\frac{110}{6}}_{6}\underbrace{\frac{001}{1}}_{1}\underbrace{\frac{101}{5}}_{5}\underbrace{\frac{011}{3}}_{3}\underbrace{\frac{111}{7}}_{7}\underbrace{\frac{100}{4}}_{4}\underbrace{\frac{000}{0}}_{0}\underbrace{\frac{110}{6}}_{6}\right)_{2} = (26153.7406)_{8}$$

$$\left(\underbrace{\frac{10}{2}}_{2}\underbrace{\frac{1100}{C}}_{6}\underbrace{\frac{0110}{6}}_{6}\underbrace{\frac{1011}{B}}_{B}.\underbrace{\frac{1111}{F}}_{F}\underbrace{\frac{0000}{0}}_{0}\underbrace{\frac{0110}{6}}_{6}\right)_{2} = (2C6B.F06)_{16}$$

# ADDITION OF UNSIGNED NUMBERS



#### O Addition of Unsigned Numbers O O

(a) The four possible cases

X	у	carry c	Sum s
0	0	0	0
0	1	0	1
1	0	0	1
1	1	1	0

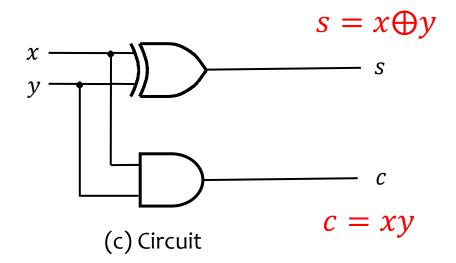
$$c = xy$$
  
 $s = x'y + xy'$   
 $= x \oplus y$   
Mod 2 addition

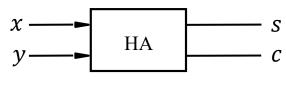
(b) Truth table

Figure 5.2. Half-adder.



#### Addition of Unsigned Numbers





(d) Graphical symbol



#### An example of addition

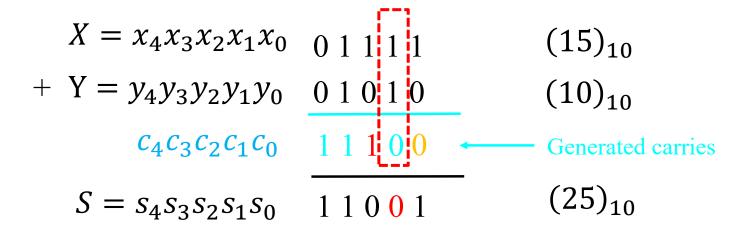


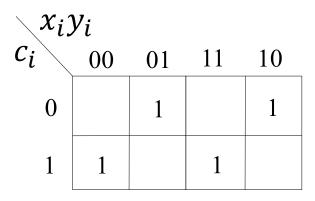
Figure 5.3. An example of addition.



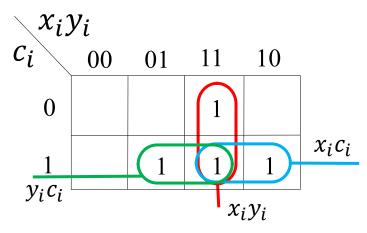
#### Full Adder

$c_i$	$x_i$	$y_i$	$c_{i+1}$	$s_i$
0	0	0	0	0
0	0	1	0	1
0	1	0	0	1
0	1	1	1	0
1	0	0	0	1
1	0	1	1	0
1	1	0	1	0
1	1	1	1	1

(a) Truth table



$$s_i = x_i \oplus y_i \oplus c_i$$



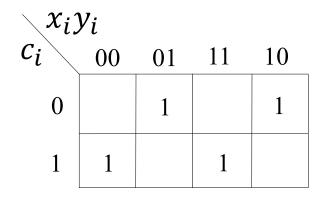
$$c_{i+1} = x_i y_i + x_i c_i + y_i c_i$$

(b) Karnaugh maps

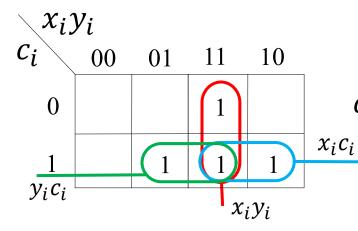
Figure 5.4. Full-adder.

#### Full Adder





$$s_i = x_i \oplus y_i \oplus c_i$$



$$s_i = c'_i x'_i y_i + c'_i x_i y'_i + c_i x'_i y'_i + c_i x_i y_i$$

$$x_i \oplus y_i = x'_i y_i + x_i y'_i \qquad \text{X-OR}$$

$$(x_i \oplus y_i)' = x'_i y'_i + x_i y_i \qquad \text{X-NOR}$$

$$s_i = c'_i(x'_iy_i + x_iy'_i) + c_i(x'_iy'_i + x_iy_i)$$
  
=  $c'_i(x_i \oplus y_i) + c_i(x_i \oplus y_i)'$   
=  $x_i \oplus y_i \oplus c_i$ 

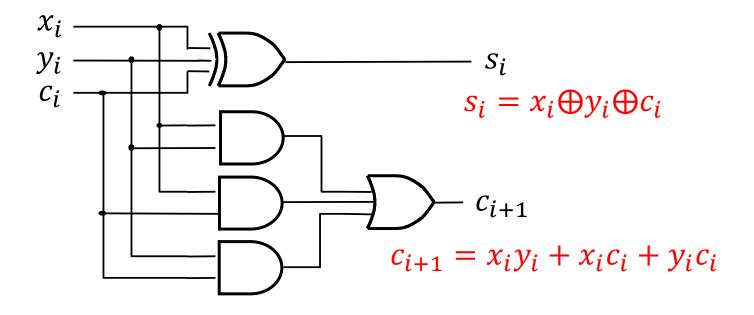
$$c_{i+1} = c'_i x_i y_i + c_i x'_i y_i + c_i x_i y'_i + c_i x_i y_i$$

$$c_{i+1} = x_i y_i + x_i c_i + y_i c_i$$



#### Full Adder





(c) Circuit

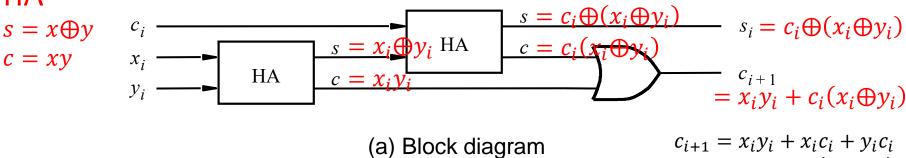
Figure 5.4. Full-adder.

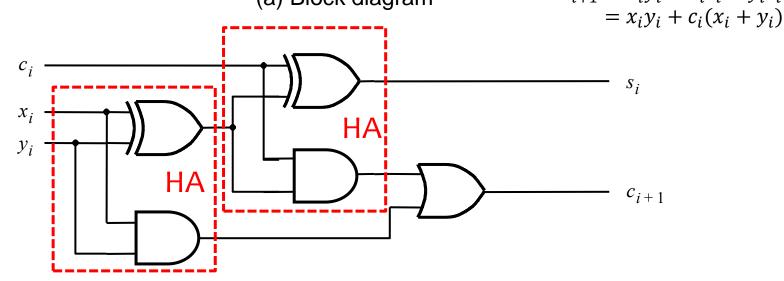


#### Decomposed Full-Adder









(b) Detailed diagram

Figure 5.5 A decomposed implementation of the full adder circuit.

### Implementation of $c_{i+1}$

$$c_{i+1} = x_i y_i + x_i c_i + y_i c_i$$

$$= x_i y_i + c_i (x_i + y_i)$$

$$c_{i+1} = x_i y_i + c_i (x_i \oplus y_i)$$

$$= x_i y_i + c_i x_i' y_i + c_i x_i y_i'$$

$$= x_i y_i + x_i c_i + y_i c_i$$

$$x_i y_i$$

$$c_i = x_i y_$$

 $c_i x_i' y_i$ 





### An example of addition

$$X = x_4 x_3 x_2 x_1 x_0 \qquad 0 \ 1 \ 1 \ 1 \qquad (15)_{10}$$

$$+ \quad Y = y_4 y_3 y_2 y_1 y_0 \qquad 0 \ 1 \ 0 \ 1 \qquad (10)_{10}$$

$$c_5 \ c_4 \ c_3 \ c_2 \ c_1 \qquad 0 \ 1 \ 1 \ 1 \ 0 \qquad (4) \ (3)^{(2)} \ (1)$$

$$S = S_4 \ S_3 \ S_2 \ S_1 \ S_0 \qquad 1 \ 1 \ 0 \ 0 \ 1 \qquad (25)_{10}$$

Figure 5.3. An example of addition.





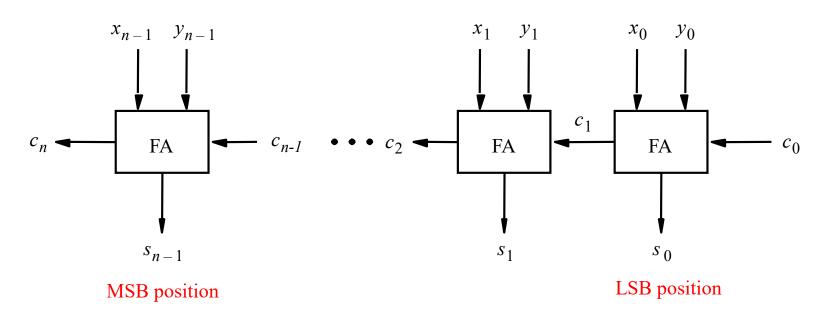
#### Ripple Carry Adder



#### The delay depends on the size of numbers

FA: full adder

Total delay  $\approx n\Delta t$ ,  $\Delta t$  is one bit delay





#### Design Example



The Logic to calculate 3A using 8 bit adder

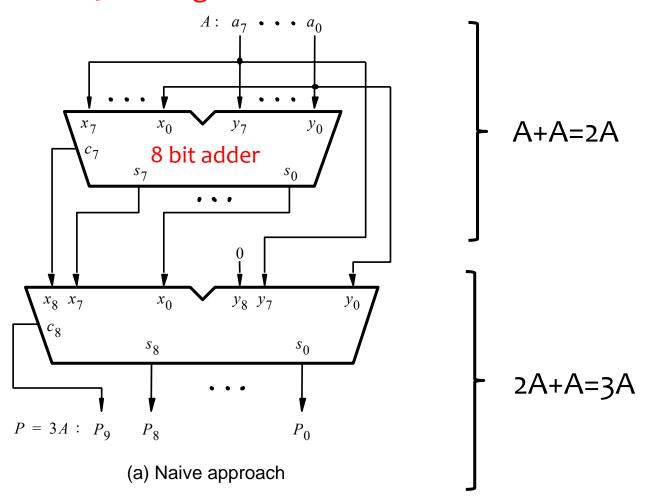
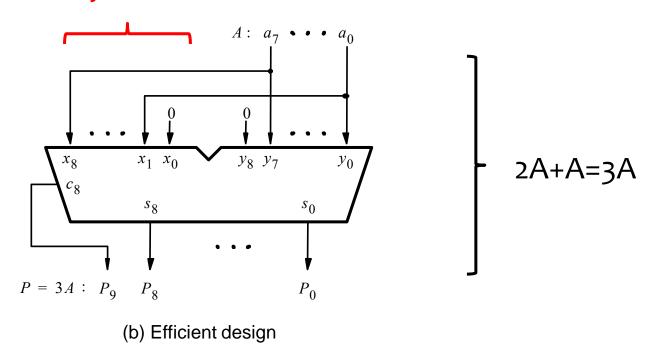


Figure 5.7. Circuit that multiplies an eight-bit unsigned number by 3

#### Design Example



#### Shift by 1 to the Left = 2A

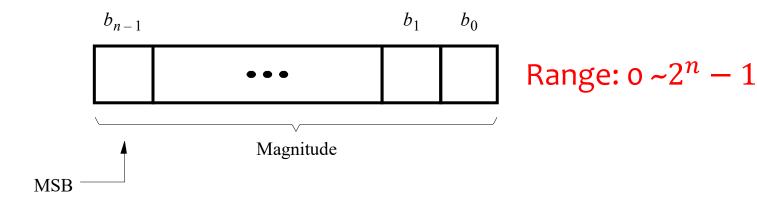


## SIGNED NUMBERS

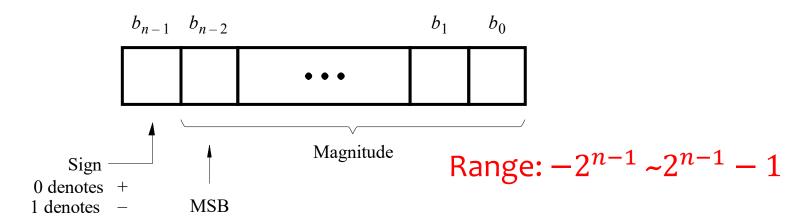


#### Signed Numbers





(a) Unsigned number



(b) Signed number



#### Signed Number



$b_3b_2b_1b_0$	Sign and magnitude	1's complement	2's complement
0111 0110 0101 0100 0011 0010 0001 0000 1000 1011 1100 1101 1110	+7 +6 +5 +4 +3 +2 +1 +0 =0 -1 -2 -3 -4 -5 -6	+7 +6 +5 +4 +3 +2 +1 +0 -7 -6 -5 -4 -3 -2 -1	+7 +6 Sign and magnitude +5: $-(2^{n-1}-1) \sim 2^{n-1} - 1$ +4 2's complement +3: $-2^{n-1} \sim 2^{n-1} - 1$ +2 +1 1's complement +0: $+0 = 0000$ -8 $-0 = 1111$ -7 -6 -5 -4 -3 -2
1111	<del>-</del> 7	-0	<del>-</del> 1

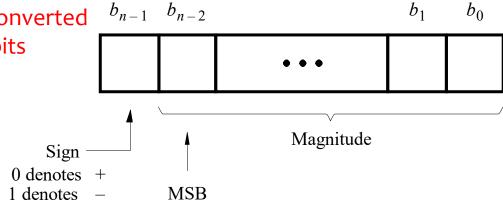


#### Sign and Magnitude

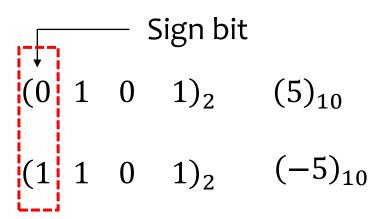


N: number to be converted  $b_{n-1}$   $b_{n-2}$ 

n: the number of bits



When n= 4





#### 1's Complement



 $b_1'$ 

 $b_0'$ 

f N

$$1's complement of N = -N$$

1's complement of 
$$N = (2^n - 1) - N$$

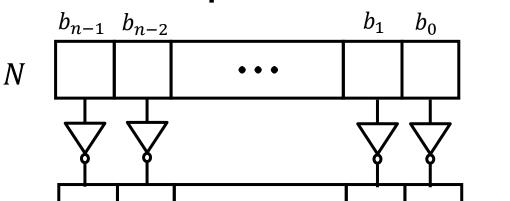
When 
$$n=4$$
  $(2^4-1) = (1 \ 1 \ 1 \ 1)_2 (15)_{10}$   $- N = (0 \ 1 \ 0 \ 1)_2 (5)_{10}$ 

 $b'_{n-1}$   $b'_{n-2}$ 

$$(1 \ 0 \ 1 \ 0)_2 \ (-5)_{10}$$



#### 2's Complement



 $b_1'$ 

 $b_0'$ 

2's complement of N = -N

2's complement of 
$$N = 2^n - N = ((2^n - 1) - N) + 1$$
  
When  $n = 4$ 

$$= 1's complement of  $N + 1$ 

$$(2^4 - 1) = (1 \ 1 \ 1 \ 1)_2 \quad (15)_{10}$$

$$- N = (0 \ 1 \ 0 \ 1)_2 \quad (5)_{10}$$

$$+ \quad (0 \ 0 \ 0 \ 1)_2$$$$

 $b'_{n-1}$   $b'_{n-2}$ 

 $(1 \ 0 \ 1 \ 1)_2 \ (-5)_{10}$ 

#### 1's complement addition



$$r = 2 \qquad r : Base \\ N : number to be converted \\ n : the number of bits$$

$$-N = (2^4 - 1) - N \qquad M + N \le (2^4 - 1)$$

$$M + N \qquad positive \qquad -M + N = -(M - N) \qquad negative$$

$$\frac{(+5)}{(+7)} \qquad \frac{0 \ 1 \ 0 \ 1}{0 \ 1 \ 1 \ 1} \qquad \frac{(-5)}{(-3)} \qquad \frac{1 \ 0 \ 1 \ 0}{1 \ 1 \ 0 \ 0}$$

$$\frac{+(+2)}{(+7)} \qquad \frac{+0 \ 0 \ 1 \ 0}{0 \ 1 \ 1 \ 1} \qquad \frac{-M + N}{(-3)} \qquad \frac{\text{No carry }!}{=(2^4 - 1) - M + N}$$

 $= (2^4 - 1) - (M - N)$ 



### 1's complement addition



$$M-N$$
 positive  $(+5)$   $0 1 0 1$   $+(-2)$   $+ 1 1 0 1$   $1 0 0 1 0$   $1 0 0 1 1$  erated!

$$-M-N=-(M+N)$$
 negative

$$\frac{(-5)}{+(-2)} + \frac{1}{1} \cdot \frac{0}{1} \cdot \frac{1}{1}$$
Carry generated! 1 0 0 0

#### Carry generated!

$$M - N$$
  
=  $M + (2^4 - 1) - N$   
=  $M - N + (2^4 - 1)$ 

$$-M - N$$

$$= (2^{4} - 1) - M + (2^{4} - 1) - N$$

$$= (2^{4} - 1) - (M + N) + (2^{4} - 1)$$

Figure 5.9. Examples of 1's complement addition.



### 2's complement addition

Figure 5.10. Examples of 2's complement addition.



Figure 5.10. Examples of 2's complement addition.



#### 2's complement subtraction

Figure 5.11. Examples of 2's complement subtraction.

### 2's complement subtraction

$$\frac{(+5)}{-(+2)} - \frac{0101}{-0010} \longrightarrow \frac{0101}{10011} \xrightarrow{\frac{M}{+(2^4 - N)}} = \frac{M - N + 2^4}{10011}$$

$$\frac{(-5)}{-(+2)} - \frac{1011}{-0010} \longrightarrow \frac{1011}{11001} \xrightarrow{\frac{4 - M}{+(2^4 - N)}} = \frac{1011}{11001} \xrightarrow{\frac{4 - M}{+(2^4 - N)}} = \frac{1011}{11001}$$
ignore

## 4 bit 2's complement $: -8 \sim +7$

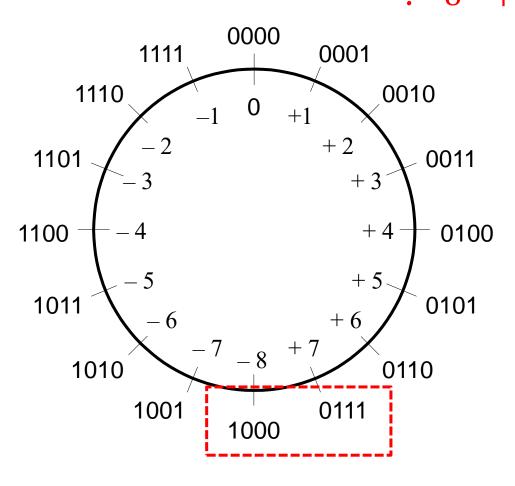


Figure 5.12. Graphical interpretation of four-bit 2's complement numbers.



#### Adder/Subtractor Unit



2's complement = 1's complement + 1
1's complement : bit inversion (NOT gate)

$$y_i \longrightarrow f = y_i \oplus 0 = y_i$$

$$y_i \longrightarrow f = y_i \oplus 1 = y_i'$$

o: addition

1: subtraction

$$\frac{y_i}{\overline{Add}/Sub} \longrightarrow f = y_i \oplus \overline{Add}/Sub = \begin{cases} y_i & if \overline{Add}/Sub = 0 \\ y_i' & if \overline{Add}/Sub = 1 \end{cases}$$



#### Adder/Subtractor Unit



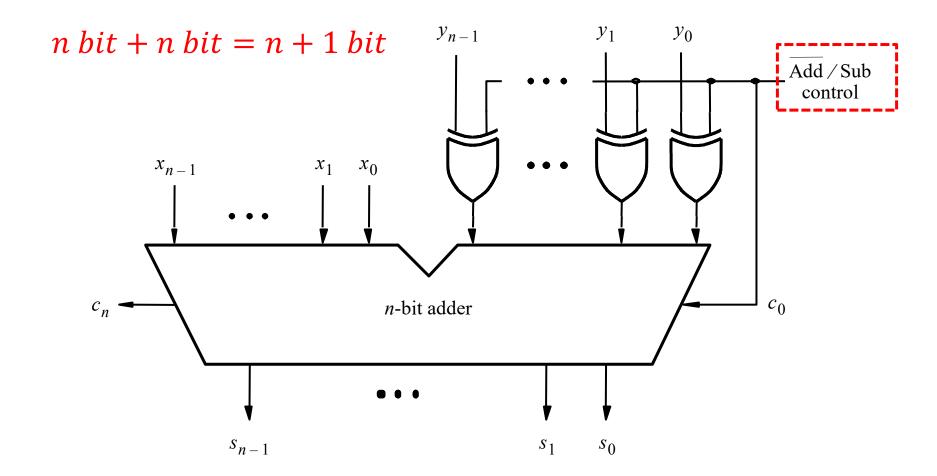


Figure 5.13. Adder/subtractor unit.



#### Arithmetic Overflow



Range of 4 bit 2'c complement binary number

$$:-2^3 \sim 2^3 - 1 = -8 \sim +7$$
 Overflow =  $c_{n-1} \oplus c_n$ 

Overflow: Out of range

Figure 5.14. Examples of determination of overflow.

#### Performance Issues

- When buying a digital system, the buyer pays particular attention to the performance and the cost.
- Superior performance comes at a higher cost.
- In n-bit ripple carry adder, the delay is approximately  $n\Delta t$ , where  $\Delta t$  is a one bit delay.
- ▶ To speed up the computer performance, it is important to find faster circuits to perform addition.



## FAST ADDER



#### Ripple-Carry Adder



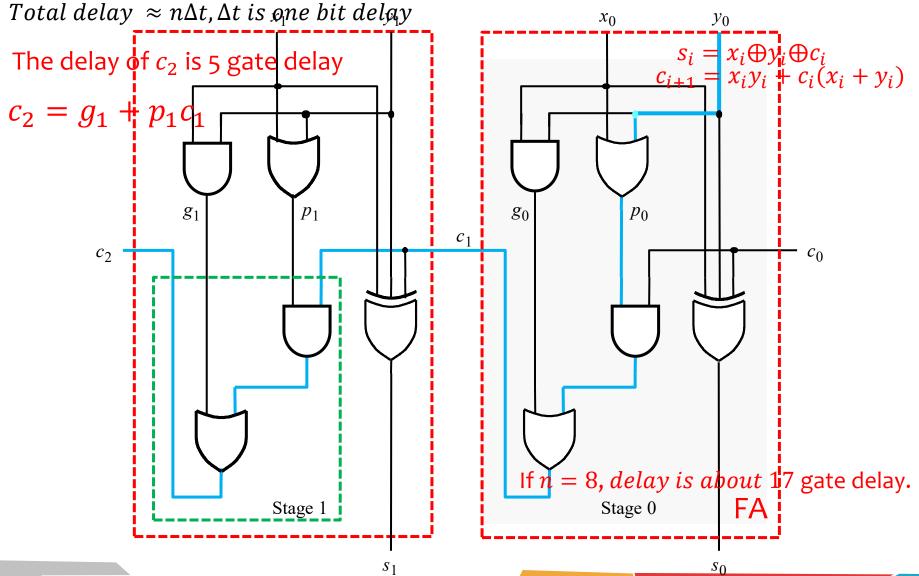


Figure 5.15. A ripple-carry adder based on expression 5.3.

#### Carry-Lookahead Adder

 $c_{i+1} = x_i y_i + x_i c_i + y_i c_i$  $c_{i+1} = x_i y_i + c_i (x_i + y_i)$ 

$$c_{i+1} = g_i + p_i c_i$$

 $c_{i+1} = g_i + p_i c_i$  Ripple Carry Adder

 $g_i = x_i y_i$ : generate function 1 x 2-input AND  $p_i = x_i + y_i$ : propagate function 1 x 2-input OR  $c_{i+1} = g_i + p_i(g_{i-1} + p_{i-1}c_{i-1})$  $c_{i+1} = g_i + p_i g_{i-1} + p_i p_{i-1} c_{i-1}$ 

$$c_{i+1} = g_i + p_i g_{i-1} + \\ p_i p_{i-1} g_{i-2} + \dots + p_i p_{i-1} \dots p_2 p_1 g_0 + p_i p_{i-1} \dots p_1 p_0 c_0$$

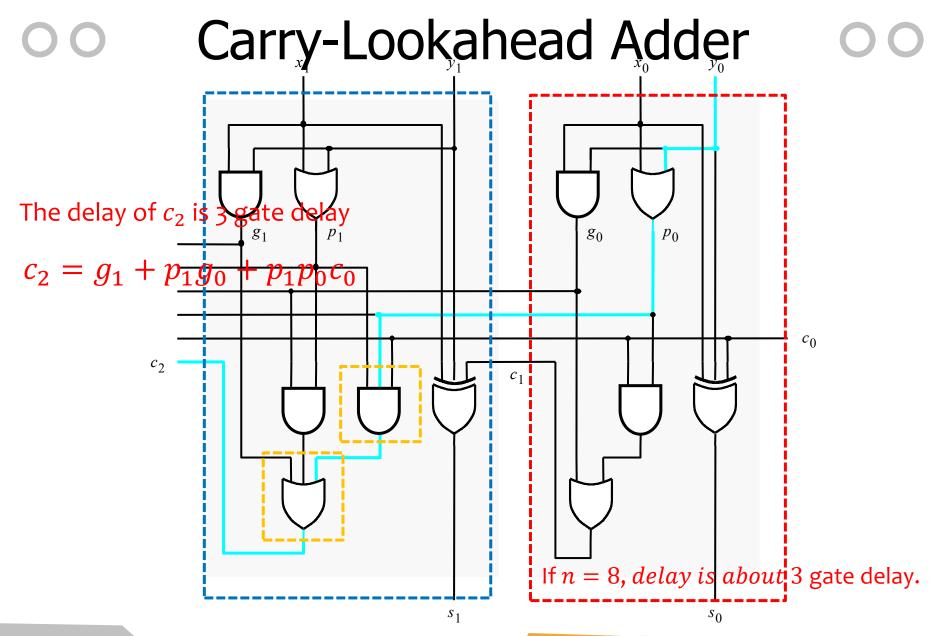


Figure 5.16. The first two stages of a carry-lookahead adder.

#### Carry-Lookahead Adder

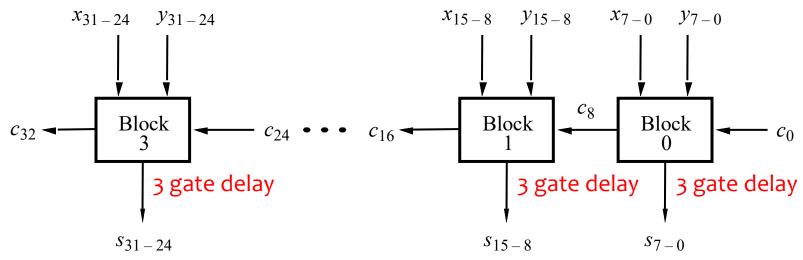
$$c_{i+1} = g_i + p_i c_i$$
 
$$c_1 = g_0 + p_0 c_0$$
 
$$c_2 = g_1 + p_1 c_1$$
 Ripple carry 
$$c_2 = g_1 + p_1 g_0 + p_1 p_0 c_0$$
 Carry Lookahead 
$$c_3 = g_2 + p_2 g_1 + p_2 p_1 g_0 + p_2 p_1 p_0 c_0$$
 
$$c_{i+1} = g_i + p_i g_{i-1} + p_i p_{i-1} g_{i-2} + \dots + p_i p_{i-1} \dots p_2 p_1 g_0 + p_i p_{i-1} \dots p_1 p_0 c_0$$

- The complexity of an n-bit-carry-lookahead adder increases rapidly as n becomes larger.
- > To reduce the complexity, a hierarchical approach can be used.

#### hierarchical carry-lookahead adder

Carry-lookahead inside block and ripple carry between blocks

Total 12 gate delay to complete  $c_{32}$ 



The carry-out signals from four blocks:  $c_8$ ,  $c_{16}$ ,  $c_{24}$ ,  $c_{32}$ 

Figure 5.17. A hierarchical carry-lookahead adder with ripple-carry between blocks

#### Second-level Lookahead

$$g_i = x_i y_i$$
$$p_i = x_i + y_i$$

$$c_{8} = g_{7} + p_{7}g_{6} + p_{7}p_{6}g_{5} + p_{7}p_{6}p_{5}g_{4} + p_{7}p_{6}p_{5}p_{4}g_{3} + p_{7}p_{6}p_{5}p_{4}p_{3}g_{2} + p_{7}p_{6}p_{5}p_{4}p_{3}p_{2}g_{1} + p_{7}p_{6}p_{5}p_{4}p_{3}p_{2}p_{1}g_{0} + p_{7}p_{6}p_{5}p_{4}p_{3}p_{2}p_{1}p_{0}c_{0}$$

$$P_{0} = p_{7}p_{6}p_{5}p_{4}p_{3}p_{2}p_{1}p_{0}$$

$$G_{0} = g_{7} + p_{7}g_{6} + p_{7}p_{6}g_{5} + \dots + p_{7}p_{6}p_{5}p_{4}p_{3}p_{2}p_{1}g_{0}$$

$$c_{8} = G_{0} + P_{0}c_{0} \quad \text{(Block 0)}$$

$$c_{16} = G_{1} + P_{1}c_{8} \quad \text{(Block 1)} \quad \text{Ripple carry between blocks}$$

$$= G_{1} + P_{1}G_{0} + P_{1}P_{0}c_{0} \quad \text{Second-level lookahead}$$

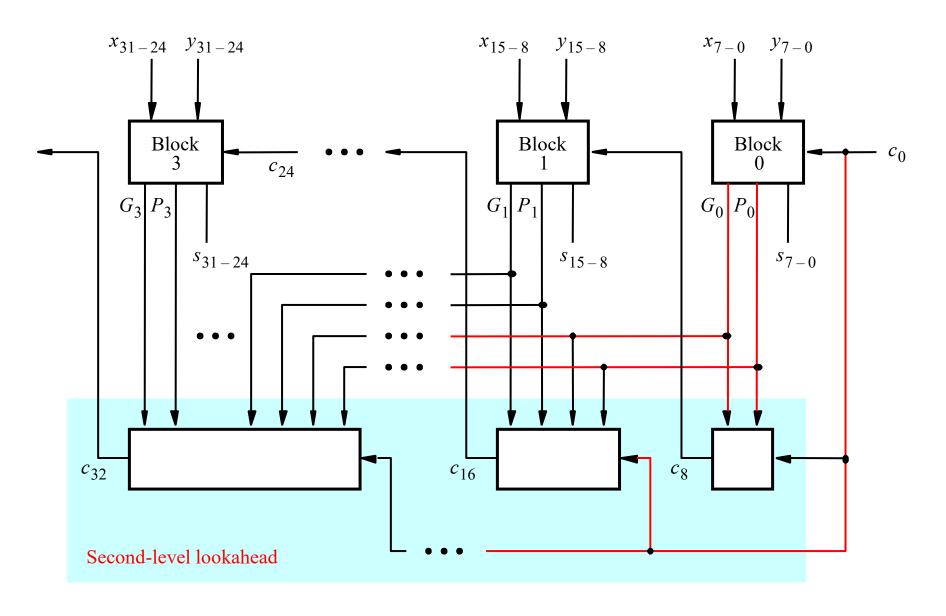


Figure 5.18. A hierarchical carry-lookahead adder.

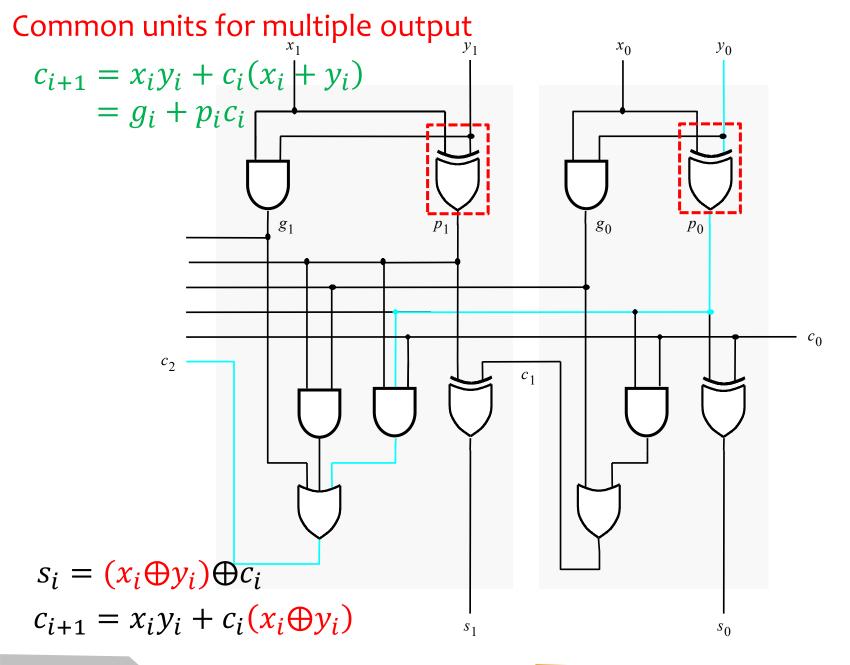


Figure 5.19. An alternative design for a carry-lookahead adder.

### O Technology Considerations

Standard SOP: two level implementation

$$c_1 = g_0 + p_0 c_0$$

$$c_2 = g_1 + p_1 g_0 + p_1 p_0 c_0$$

$$\vdots$$

$$c_8 = g_7 + p_7 g_6 + p_7 p_6 g_5 + p_7 p_6 p_5 g_4 + p_7 p_6 p_5 p_4 g_3 + p_7 p_6 p_5 p_4 p_3 g_2 + p_7 p_6 p_5 p_4 p_3 p_2 g_1 + p_7 p_6 p_5 p_4 p_3 p_2 p_1 g_0 + p_7 p_6 p_5 p_4 p_3 p_2 p_1 p_0 c_0$$

2-input AND, 3-input AND, 4-input AND, 5-input AND, 6-input AND, 7-input AND, 9-input AND, 9-input OR

Cost = 
$$9 + (2+3+4+5+6+7+8+9+9)=62$$

Suppose that the maximum fan-in of the gates is four inputs.

#### Multilevel implementation

$$c_8 = (g_7 + p_7 g_6 + p_7 p_6 g_5 + p_7 p_6 p_5 g_4) + ((p_7 p_6 p_5 p_4)(g_3 + p_3 g_2 + p_3 p_2 g_1 + p_3 p_2 p_1 g_0) + (p_7 p_6 p_5 p_4)(p_3 p_2 p_1 p_0)c_0$$

3 x 2-input AND, 3 x 3-input AND, 5 x 4-input AND, 2 x 3-input OR, 2 x 4-input OR

Cost =15 + (2x3+3x3+5x4+2x3+2x4)=64

## O O O MULTIPLICATION



#### Multiplication



```
n \ bit + n \ bit = n + 1 \ bit
n \ bit \times n \ bit = 2n \ bit
```

Multiplicand M Multiplier Q	(14) (11)	1 1 1 0 x 1 0 1 1	M	$m_3 m_2 m_1 m_0$
		$ \begin{array}{c} 1 \ 1 \ 1 \ 0 \\ 0 \ 0 \ 0 \ 0 \end{array} $	$\times$ Q	$\times q_3 q_2 q_1 q_0$
Product P	(154)	$\frac{1110}{10011010}$	P	$p_7 p_6 p_5 p_4 p_3 p_2 p_1 p_0$

(a) Multiplication by hand

Figure 5.31. Multiplication of unsigned numbers.



#### Multiplication



```
      Multiplicand M
      (11)
      1110

      Multiplier Q
      (14)
      x 1011

      Partial product 0
      1110

      Partial product 1
      10101

      Partial product 2
      01010

      Partial product 2
      01010

      Product P
      (154)
      1001101
```

(b) Multiplication for implementation in hardware



#### Multiplication

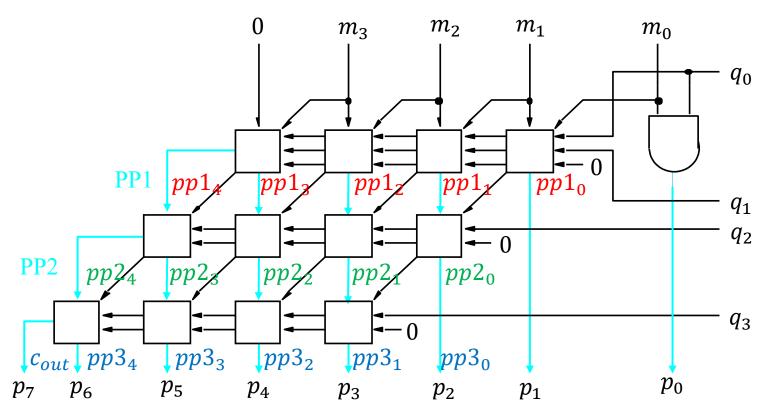


```
PP0 = m_3 q_0 \quad m_2 q_0 \quad m_1 q_0 \quad m_0 q_0
PP0:
                     pp0_3 pp0_2 pp0_1 pp0_0
           + m_3 q_1 m_2 q_1 m_1 q_1 m_0 q_1
                pp1_4 pp1_3 pp1_2 pp1_1 pp1_0 p_0
PP1:
PP1:
                  0 pp1_4 pp1_3 pp1_2 pp1_1
                m_3q_2 m_2q_2 m_1q_2 m_0q_2
                 pp2_4 pp2_3 pp2_2 pp2_1 pp2_0
PP2:
PP2:
                  0 pp2_4 pp2_3 pp2_2 pp2_1
                m_3q_3 m_2q_3 m_1q_3 m_0q_3
PP3:
            c_{out} pp3_4 pp3_3 pp3_2 pp3_1 pp3_0
  P:
              p_7 p_6 p_5 p_4 p_3 p_2
```

$$pp0_0 = pp1_0 = p_0$$
  
 $pp1_1 = pp2_0 = p_1$ 

$$pp2_1 = pp3_0 = p_2$$

#### 4 x 4 multiplier circuit



(a) Structure of the circuit

Figure 5.32. A 4 x 4 multiplier circuit.

#### 4 x 4 multiplier circuit



Module for PP1

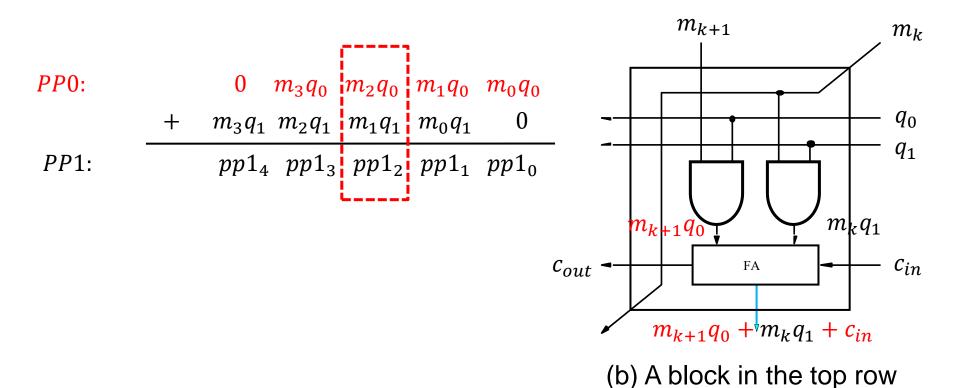


Figure 5.32. A 4 x 4 multiplier circuit.

#### oo 4 x 4 multiplier circuit



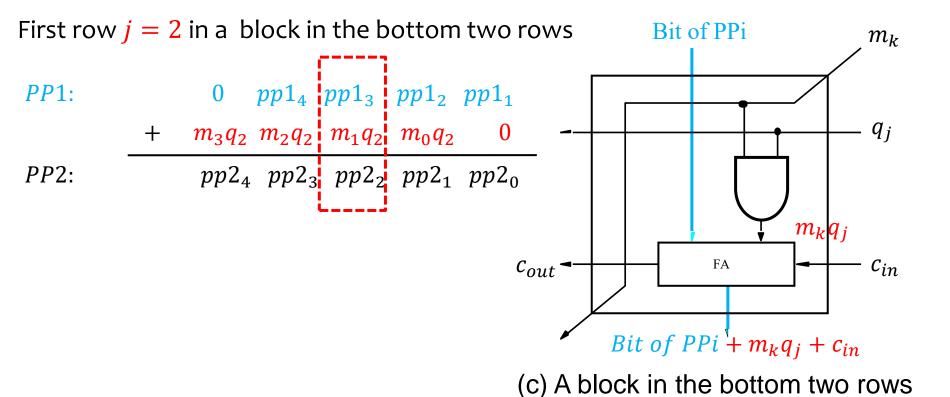


Figure 5.32. A 4 x 4 multiplier circuit.

Module for PP2 and P

#### Multiplication of signed numbers

$$\begin{array}{c} 0 & 1 & 1 & 1 & 0 \\ x & 0 & 1 & 0 & 1 & 1 \\ \hline & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ & + & 0 & 0 & 1 & 1 & 1 & 0 \\ \hline & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ & + & 0 & 0 & 1 & 0 & 1 & 0 \\ & + & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ & + & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ & + & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ \end{array}$$

(a) Positive multiplicand

Figure 5.33. Multiplication of signed numbers.

#### Multiplication of signed numbers



$$(-14)$$
 (+11)

Partial product 0

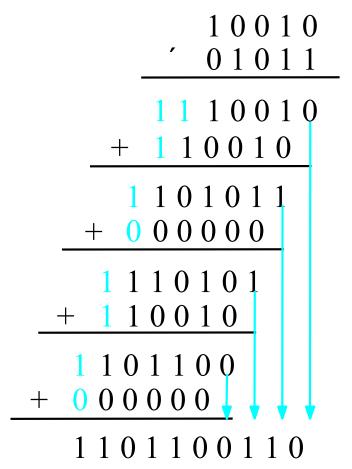
Partial product 1

Partial product 2

Partial product 3

Product P

$$(-154)$$



(b) Negative multiplicand

Figure 5.33. Multiplication of signed numbers.

# OOO OTHER NUMBER REPRESENTATIONS

#### **Fixed-Point Numbers**

$$B = b_{n-1}b_{n-2}\cdots b_1b_0 \cdot b_{-1}b_{-2}\cdots b_{-k}$$

$$V(B) = \sum_{i=-k}^{n-1} b_i \times 2^i$$

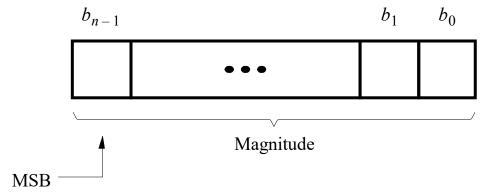
The position of radix point is assumed; hence the name fixed-point number.

If the radix point is not shown, then it is assumed to be to the right of the least-significant digit, which means that the number is an integer.

#### 00

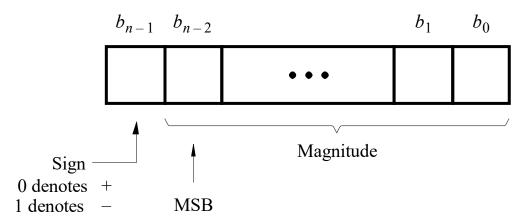
#### Fixed-Point Numbers





Range:  $0 \sim 2^n - 1$ 

(a) Unsigned number



Range:  $-2^{n-1} \sim 2^{n-1} - 1$ 

(b) Signed number

#### Floating-Point Numbers



Using excess-127 is convenient for adding and subtracting floating-point number

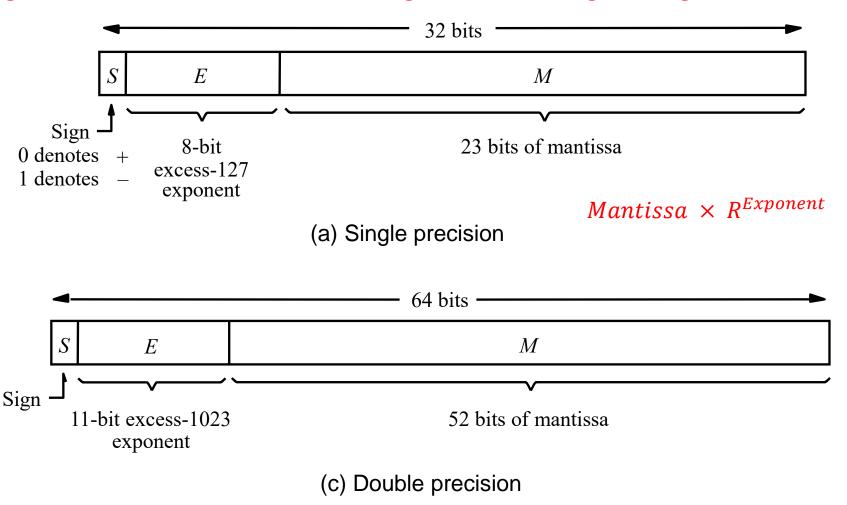


Figure 5.34. IEEE Standard floating-point formats.

#### OOO Floating-Point Numbers OOO

- Single-Precision Floating-Point Format
  - ▶ Exponent = E 127
    - ▶ Range: 0~255, 0: real "0", 255: "∞"
  - $Value = \pm 1.M \times 2^{E-127}$
  - The exponent range of  $2^{-126}$  to  $2^{127}$  corresponds to about  $10^{\pm 38}$
- Double-Precision Floating-Point Format
  - ▶ Exponent = E 1023
  - $Value = \pm 1.M \times 2^{E-1023}$
  - The exponent range of  $2^{-1022}$  to  $2^{1023}$  corresponds to about  $10^{\pm 308}$

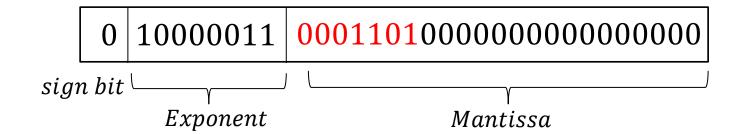
#### Floating-Point Numbers 000

- Single-Precision Floating-Point Format
  - $Value = +1.M \times 2^{E-127}$

$$(17.625)_{10} = (10001.101)_2$$

$$10001.101 = 1.0001101 \times 2^4$$

$$E = 4 + 127 = 10000011 = (131)_{10}$$



#### Floating-Point Numbers 000

- Single-Precision Floating-Point Format
  - $Value = +1.M \times 2^{E-127}$
  - ▶ Exponent = E 127
  - ▶ Range of Exponent:  $-2^7 \sim 2^7 1 = -128 \sim 127$
  - Range of E: Exponent + 127 =  $-1 \sim 254$  =  $0 \sim 255$
  - $(-1)_{10} = (111111111)_2 = (255)_{10}$

#### Binary-Coded-Decimal Representation

	Decimal digit	BCD code	
	0	0000	$(13)_{10} = (1101)_2 = (D)_{16}$
	1	0001	
	2	0010	$=(0001)_2(0011)_2$
	3	0011	
	4	0100	
	5	0101	<u>;</u> !
Don't care term	6	0110	
1010, 1011, 1100,	7	0111	
1101, 1110, 1111	8	1000	
	9	1001	$(0001)_2 (0011)_2$

Table 5.3. Binary-coded decimal digits.

#### Addition of BCD digits

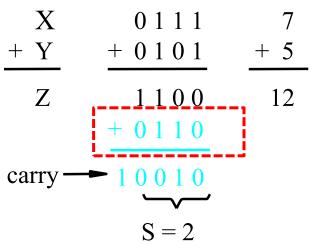
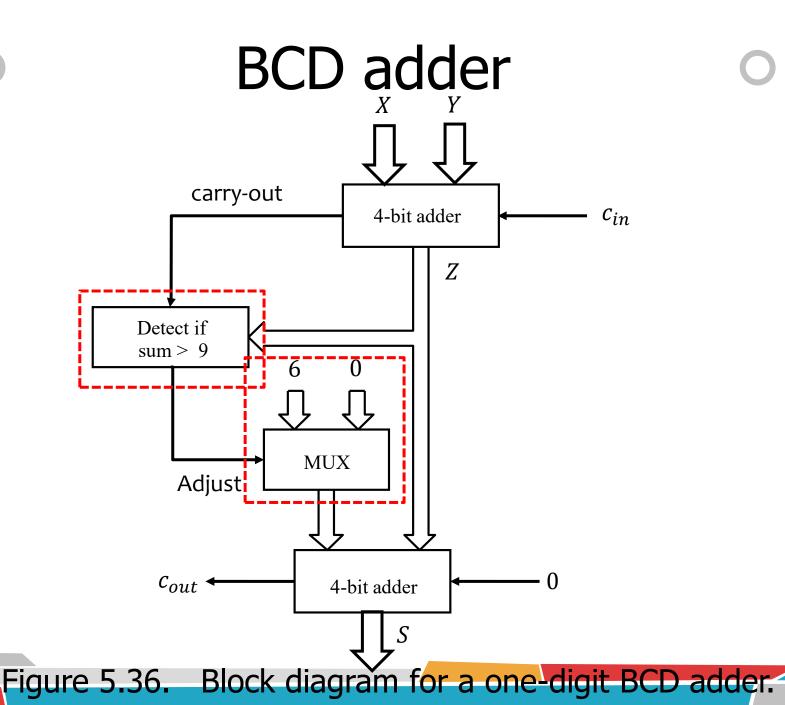


Figure 5.35. Addition of BCD digits.

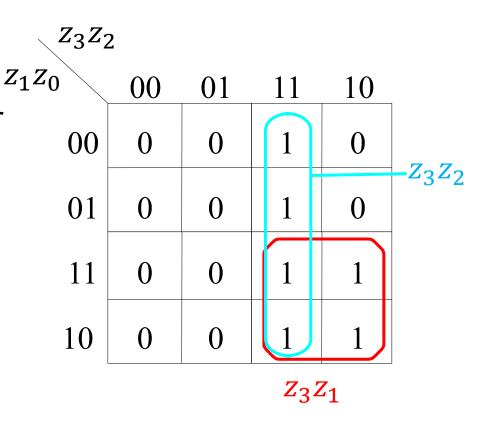


#### 00

#### Detect if sum > 9

00

- (1) If carry out of first 4-bit adder is occurred, i.e.  $c_4=1$
- (2) If sum of first 4-bit adder Z > 1,  $z_3 z_2 z_1 z_0 > 1001$



$$c_{out} = c_4 + z_3 z_2 + z_3 z_1$$

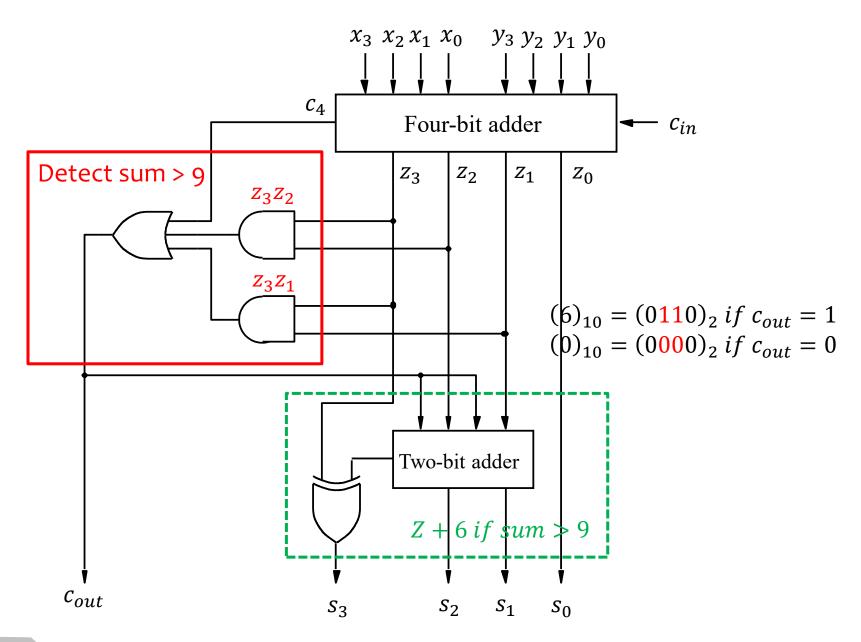


Figure 5.39. Circuit for a one-digit BCD adder.



#### **Gray Code**



Table 1-6 Gray Code

Gray code	Decimal equivalent
-	
0000	0
0001	1
0011	2
0010	3
0110	4
0111	5
0101	6
0100	7
1100	8
1101	9
1111	10
1110	11
1010	12
1011	13
1001	14
1000	15

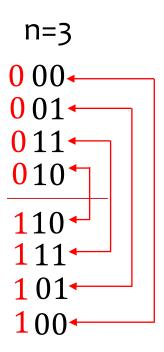
Successive Gray code has 1 bit difference only

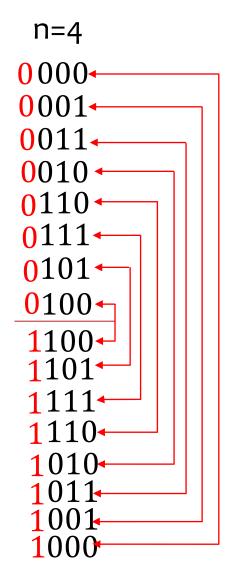
#### O Generation of Gray Code



n=1	n=2
0	00
1	01
	11
	10←

RBC (Reflected Binary Code)

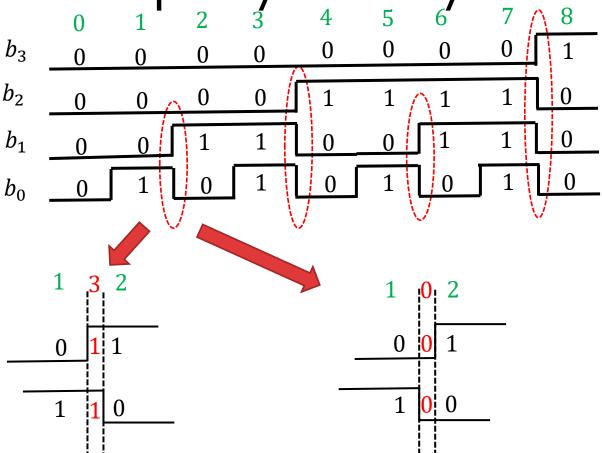




00

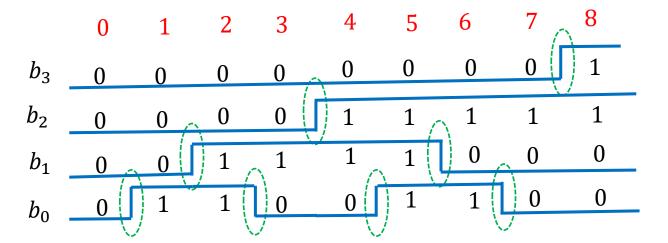
Property of Gray Code





**Binary Counter** 

## Property of Gray Code



## ASCII CHARACTER CODE

-	Bit positions	Bit positions 654								
	3210	000	001	010	011	100	101	110	111	
-	0000	NUL	DLE	SPAC	E 0	@	P		p	
128 characters English only	0001	$\mathbf{SOH}$	DC1	!	1	$\mathbf{A}$	$\mathbf{Q}$	$\mathbf{a}$	$\mathbf{q}$	
	0010	$\mathbf{S}\mathbf{T}\mathbf{X}$	DC2	"	<b>2</b>	$\mathbf{B}$	$\mathbf{R}$	b	r	
	0011	$\mathbf{E}\mathbf{T}\mathbf{X}$	DC3	#	3	$\mathbf{C}$	$\mathbf{s}$	$\mathbf{c}$	s	
	0100	EOT	DC4	\$	4	D	${f T}$	d	t	
	0101	ENQ	NAK	%	5	${f E}$	$\mathbf{U}$	e	u	
	0110	$\mathbf{ACK}$	SYN	&	6	$\mathbf{F}$	V	f	$\mathbf{v}$	
	0111	$\operatorname{BEL}$	ETB	,	7	$\mathbf{G}$	W	g	w	
	1000	$\mathbf{BS}$	CAN	(	8	$\mathbf{H}$	$\mathbf{X}$	$\mathbf{h}$	$\mathbf{x}$	
	1001	$\mathbf{HT}$	$\mathbf{E}\mathbf{M}$	)	9	Ι	$\mathbf{Y}$	i	y	
	1010	LF	SUB	*	:	J	${f z}$	j	$\mathbf{z}$	
	1011	VT	$\mathbf{ESC}$	+	;	$\mathbf{K}$	[	$\mathbf{k}$	{	
	1100	FF	FS	,	<	${f L}$	\	1		
	1101	$\mathbf{C}\mathbf{R}$	$\mathbf{G}\mathbf{S}$	-	_	$\mathbf{M}$	]	$\mathbf{m}$	}	
	1110	so	$\mathbf{RS}$	•	>	N	^	$\mathbf{n}$	~	
_	1111	SI	US	/	?	0		0	DEL	
	NUL	Null/Idle		S	SI		Shift i	in		
	SOH	Start of header Start of text		I	$rac{ ext{DLE}}{ ext{DC1-DC4}}$		Data link escape Device control Negative acknowledgement Synchronous idle End of transmitted block Cancel (error in data) End of medium			
	STX			I						
	$\mathbf{E}\mathbf{T}\mathbf{X}$	End of te	End of text NAK							
	EOT	End of transmission Enquiry Acknowledgement Audible signal Back space		ion S	SYN ETB CAN EM SUB					
	$\mathbf{E}\mathbf{N}\mathbf{Q}$									
	$\mathbf{ACQ}_{\mathbf{q}}$			t (						
	$\mathbf{BEL}$			I						
	$\mathbf{BS}$						Special sequence			
	HT	Horizontal tab			ESC		Escape			
	${f LF}$	Line feed			FS		File separator			
	VT	Vertical tab Form feed Carriage return			GS RS US		Group separator			
	$\mathbf{FF}$						Record separator			
	$\mathbf{C}\mathbf{R}$						Unit separator			
	SO	Shift out			DEL			m Delete/Idle		
	Bit positions of code format = $\begin{bmatrix} 6 & 5 & 4 & 3 & 2 & 1 & 0 \end{bmatrix}$									

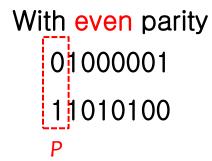
Table 5.4. The seven-bit ASCII code.

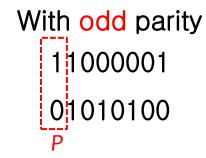
## Parity



Error-Detecting Code (parity checker)

$$b_6b_5b_4b_3b_2b_1b_0$$
  
ASCII A = 1 0 0 0 0 0 1  
ASCII T = 1 0 1 0 1 0 0





Even Parity

$$P = b_6 \oplus b_5 \oplus b_4 \oplus b_3 \oplus b_2 \oplus b_1 \oplus b_0$$

$$\bigcup$$

$$C = P \oplus b_6 \oplus b_5 \oplus b_4 \oplus b_3 \oplus b_2 \oplus b_1 \oplus b_0$$

## Parity Checker



> Even Parity

Case of Character 'A'

$$P = b_6 \oplus b_5 \oplus b_4 \oplus b_3 \oplus b_2 \oplus b_1 \oplus b_0$$
$$= 1 \oplus 0 \oplus 0 \oplus 0 \oplus 0 \oplus 0 \oplus 1 = 0$$

No Error

$$C = 0 \oplus 1 \oplus 0 \oplus 0 \oplus 0 \oplus 0 \oplus 1 = 0$$

Error occurred at  $b_4$ 

$$C = 0 \oplus 1 \oplus 0 \oplus 1 \oplus 0 \oplus 0 \oplus 0 \oplus 1 = 1$$

## O Binary Code for Text

- 000
- ASCII (American Standard Code for Information Interchange)
  - English only
  - How can we represent the other languages such as Hangeul, Chinese character, etc?
- Uni-Code
  - Unicode Consortium
  - UTF (Unicode Transformation Format): UTF-8, UTF-16, UTF-32
  - UCS (Universal Coded Character Sets): UCS-2 (16 bit), UCS-4
     (32 bit)
- EBCDIC (Extended Binary Coded Decimal Interchange Code)
  - eight-bit character encoding used mainly on IBM mainframe and IBM midrange computer operating systems

# EXAMPLES OF SOLVED PROBLEMS

## Example 5.7



#### Conversion the decimal number into a hexadecimal number

Convert (14959)<sub>10</sub>

		Remainder	Hex digit	
14959 ÷ 16 =	934	15	F	LSB
934 ÷ 16 =	58	6	6	
58 ÷ 16 =	3	10	Α	
3 ÷ 16 =	0	3	3	MSB

Result is (3A6F)<sub>16</sub>

Figure 5.40. Conversion from decimal to hexadecimal.

## Example 5.8



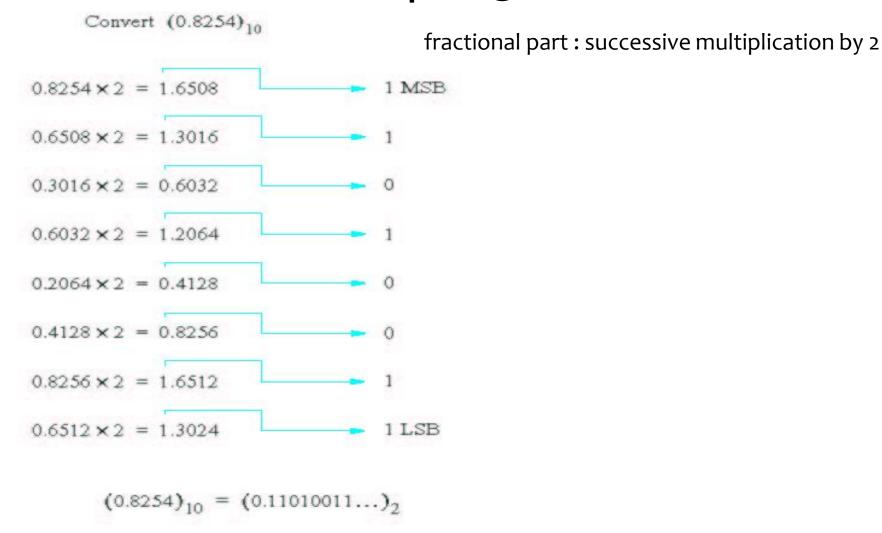


Figure 5.41. Conversion of fractions from decimal to binary.

## Example 5.9



Convert (214.45)<sub>10</sub>

$$\frac{214}{2} = 107 + \frac{0}{2} \qquad 0 \text{ LSB}$$

$$\frac{107}{2} = 53 + \frac{1}{2} \qquad 1$$

$$\frac{53}{2} = 26 + \frac{1}{2} \qquad 1$$

$$\frac{26}{2} = 13 + \frac{0}{2} \qquad 0$$

$$\frac{13}{2} = 6 + \frac{1}{2} \qquad 1$$

$$\frac{6}{2} = 3 + \frac{0}{2} \qquad 0$$

$$\frac{3}{2} = 1 + \frac{1}{2} \qquad 1$$

$$\frac{1}{2} = 0 + \frac{1}{2} \qquad 1 \text{ MSB}$$

Integer part: successive division by 2

 $0.45 \times 2 = 0.90$  0 MSB  $0.90 \times 2 = 1.80$   $0.80 \times 2 = 1.60$   $0.60 \times 2 = 1.20$   $0.20 \times 2 = 0.40$   $0.40 \times 2 = 0.80$  $0.80 \times 2 = 1.60$  1 LSB

 $(214.45)_{10} = (11010110.0111001...)_2$ 

fractional part: successive multiplication by 2

Figure 5.42. Conversion of fixed point numbers from decimal to binary.

## Example 5.10



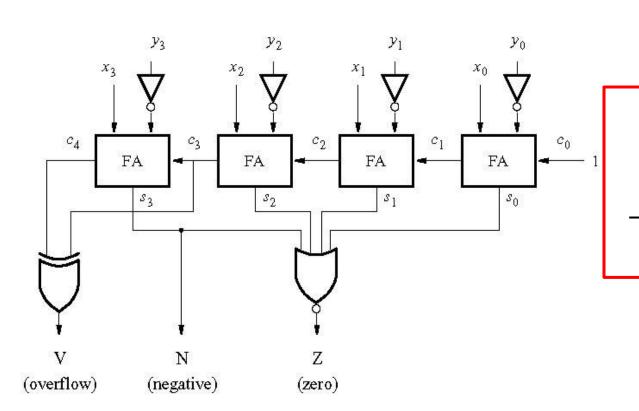


Figure 5.43. A comparator circuit.

$$X = x_3 x_2 x_1 x_0$$
$$Y = y_3 y_2 y_1 y_0$$

$$X = x_3 x_2 x_1 x_0$$

$$-Y = y_3' y_2' y_1' y_0'$$

$$+1$$

$$c_4S = c_4s_3s_2s_1s_0$$

$$S = s_3 s_2 s_1 s_0$$
$$V = c_4 \oplus c_3$$

$$N = s_3$$

$$Z = (s_3 + s_2 + s_1 + s_0)'$$
$$= \begin{cases} 1 & \text{if } X - Y = 0 \\ 0 & \text{otherwise} \end{cases}$$

## Example 5.10



$$Z = \begin{cases} 1 & if \ X = Y \\ 0 & if \ X \neq Y \end{cases}$$

$$N = \begin{cases} 1 & if \ X < Y \\ 0 & if \ X \geq Y \end{cases}$$

$$P = N'Z' = \begin{cases} 1 & if \ X > Y \\ 0 & if \ X \leq Y \end{cases}$$

## Summary

- ▶ The circuit, which implement the addition of two bits, is called a half adder.
- A full adder generates sum and carry out by using the carry in bit as well as two input bits.
- An n-bit ripple carry adder can be easily implemented by connecting full adders.
- The delay of an n-bit ripple carry adder depends on the size of the numbers, that is,  $Delay \approx n\Delta t$ ,  $\Delta t$  is the one bit gate delay.
- Subtraction can be carried out by adder and the complement.

## Summary



- Arithmetic overflow can be detected by carry out bits.
- To improve the performance of an ripple carry adder, a carry-lookahead adder can be used as a fast adder.
- Multiplication can be implemented by adder, shift operation and memory.
- Hardware multiplier accelerates the speed of multiplication.