

Week 8 - T1 2020

Arithmetic Circuits

ELEC2141: Digital Circuit Design



Overview

Arithmetic circuits

- Iterative combinational circuits

Binary adders

Unsigned binary addition

Signed binary addition

Overflow

Reading: Mano - Chapter 3: 3.8-3.11

Arithmetic circuits

Circuits that perform arithmetic operations are important part of digital design

These circuits form the basis for a microprocessor's **ALU** and commonly used in many other applications

Arithmetic circuits often designed to operate on binary input vectors and produce binary output vectors

Iterative combinational circuits

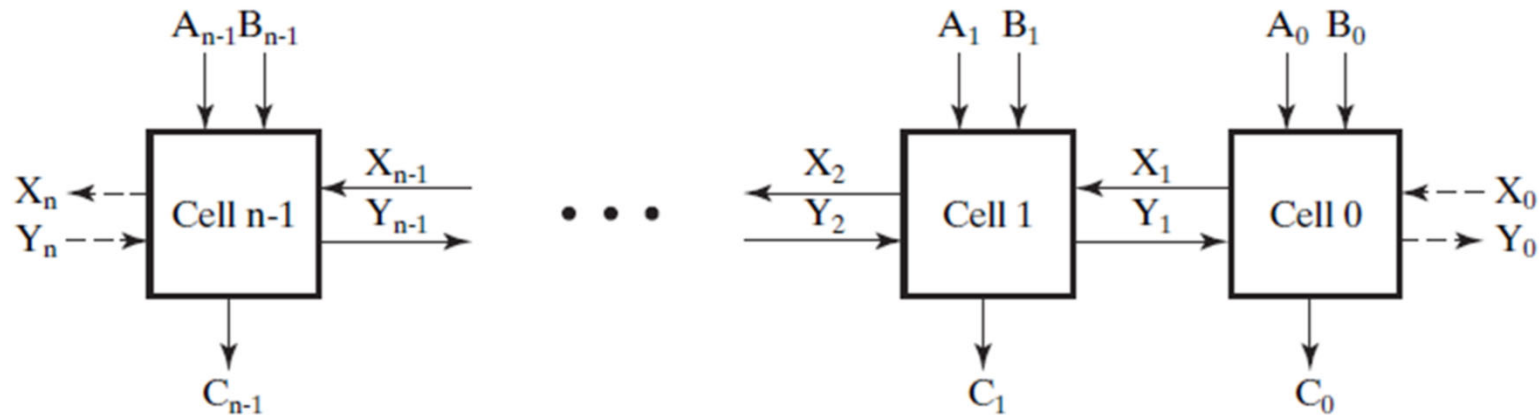
The function implemented requires that the same sub-function be applied to each bit position

Can design functional block for the sub-function and duplicate it to obtain functional block for an overall function

Each sub-function block is referred to as a *cell*

An array of interconnected cells forms an *iterative array*

Iterative combinational circuits



Example: $n = 32$ means 64 inputs and 32 outputs

2^{64} truth table rows - impractical!

Iterative array takes advantage of the regularity to make the design feasible

Binary adders

Binary addition is used frequently

Similar to decimal addition but with two possible digits: 0 and 1

Example - 8-bit addition:

$$\begin{array}{r} 77 \\ + 86 \\ \hline \end{array} \quad \begin{array}{r} 01001101 \\ + 01010110 \\ \hline \end{array}$$

Half adder

A *half adder* generates the sum of two input bits

The two inputs are represented by X and Y

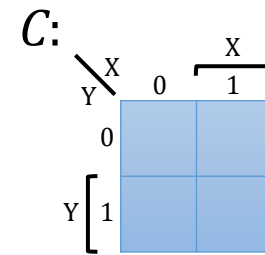
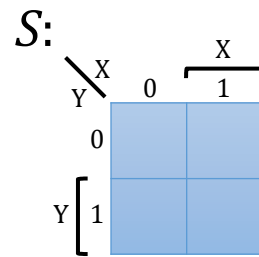
The sum is expressed as a *sum* (S) and a *carry* (C)

X	0	1	0	1
+ Y	+ 0	+ 0	+ 1	+ 1
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
C S				

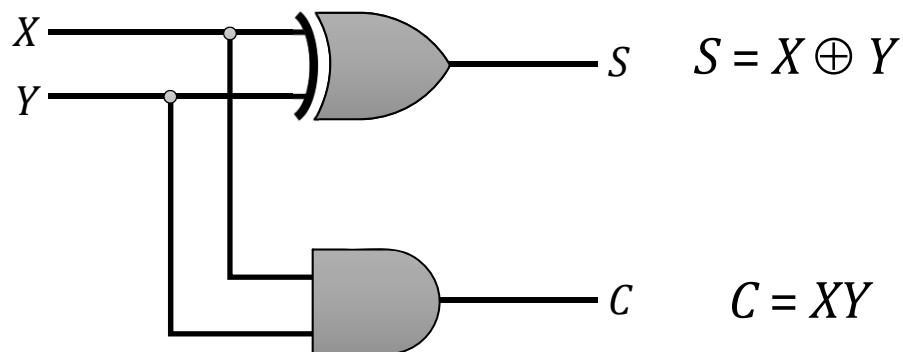
Half adder

The half adder can be specified as a truth table and the output equations can be derived

X	Y	C	S
0	0		
0	1		
1	0		
1	1		



The logic diagram for the half adder



Full adder

A full adder generates the sum of three input bits

Two inputs (X and Y) represent the two significant bits to be added

The third input (Z) represents the carry from the previous significant bit

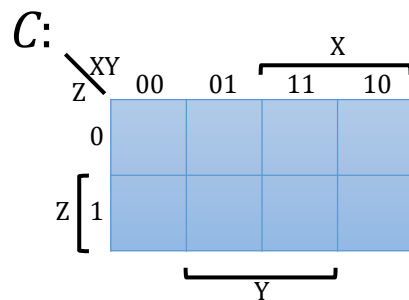
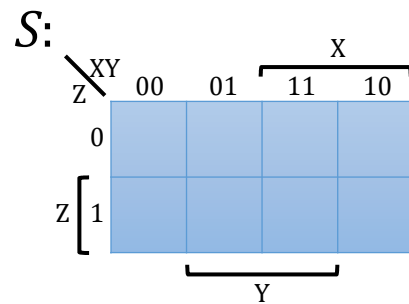
Two outputs represent the *sum* (S) and *carry* (C)

Z	0	0	0	0	1	1	1	1
X	0	0	1	1	0	0	1	1
+ Y	+ 0	+ 1	+ 0	+ 1	+ 0	+ 1	+ 0	+ 1
C S	0 0	0 1	0 1	1 0	0 1	1 0	1 0	1 1

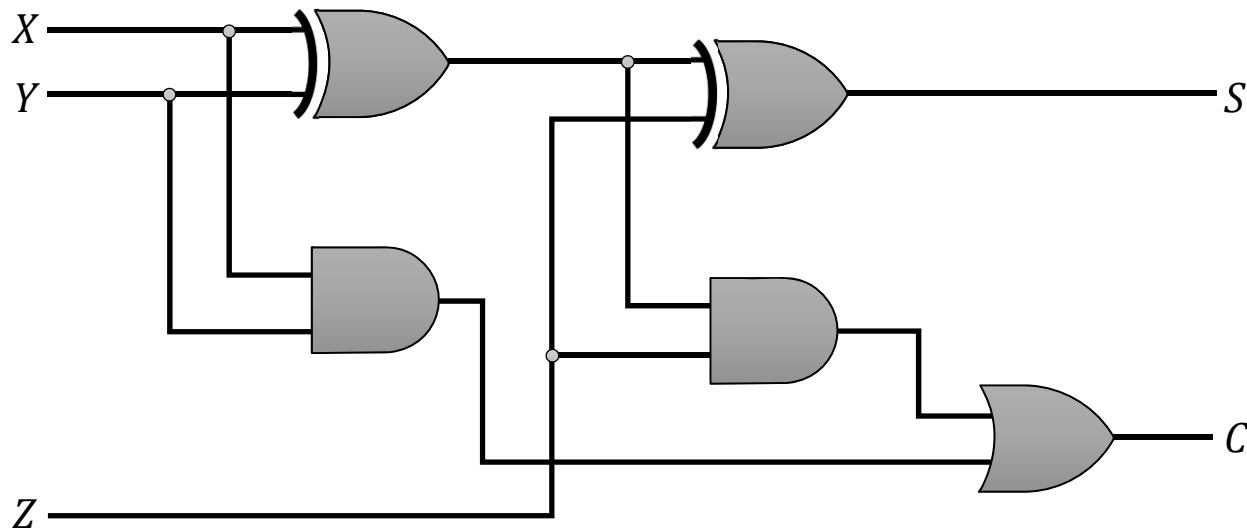
Full adder

A Full Adder can be specified as a truth table and the output equations can be derived

X	Y	Z	C	S
0	0	0		
0	0	1		
0	1	0		
0	1	1		
1	0	0		
1	0	1		
1	1	0		
1	1	1		



Full adder logic diagram



$$S = X \oplus Y \oplus Z$$
$$C = XY + Z(X \oplus Y)$$

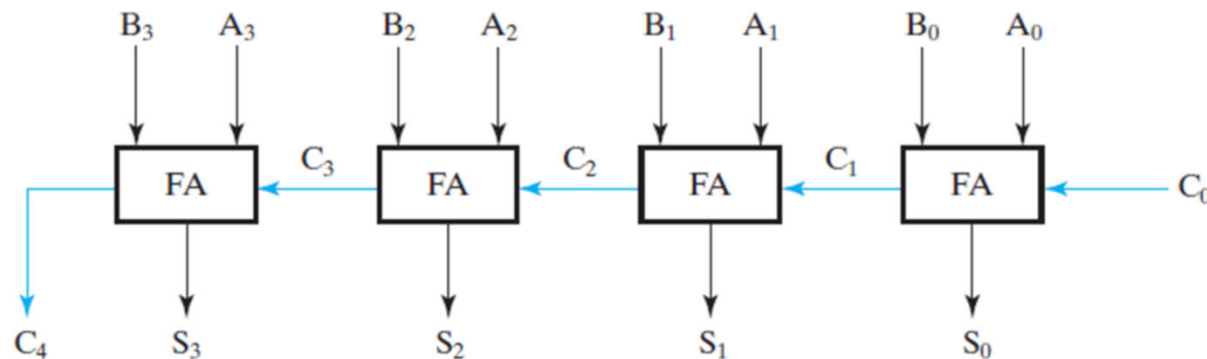
Binary ripple-carry adder

An n -bit parallel adder formed by cascading n full adders

Apply all n -bit inputs simultaneously

Connect the carry output from one full adder to the carry input of the next

The carry propagate through, from the LSB to the MSB



C_i	0	0	1	1	0
A_i		1	0	1	1
B_i		+ 0	0	1	1
S_i			1	1	1
C_{i+1}			0	0	1

Binary ripple-carry adder

Ripple-carry adders have long circuit delays

The carry has to propagate through many gates until the final result is obtained

Each of the n full-adders introduce two gate-delays in the carry path

Example: for a 16-bit adder, the delay is 32 gate delays (+ overhead)

Binary carry lookahead adder

Define: Carry *Generate*

$$G_i = A_i B_i$$

Must generate carry when $A = B = 1$

Define: Carry *Propagate*

$$P_i = A_i \oplus B_i$$

The carry-out will equal carry-in

Express the sum (S) and carry (C) in terms of generate/propagate:

$$S_i = (A_i \oplus B_i) \oplus C_i$$

$$= P_i \oplus C_i$$

$$C_{i+1} = A_i B_i + C_i (A_i \oplus B_i)$$

$$= G_i + C_i P_i$$

$$\begin{array}{rcccccc} C_i & 0 & 1 & 0 & 0 & 0 \\ A_i & & 0 & 1 & 0 & 0 \\ B_i & & + 0 & 1 & 0 & 0 \\ \hline S_i & & 1 & 0 & 0 & 0 \\ C_{i+1} & & 0 & 1 & 0 & 0 \end{array}$$

$$\begin{array}{rcccccc} C_i & 0 & 1 & 1 & 0 & 0 \\ A_i & & 0 & 0 & 1 & 0 \\ B_i & & + 0 & 1 & 1 & 0 \\ \hline S_i & & 1 & 0 & 0 & 0 \\ C_{i+1} & & 0 & 1 & 1 & 0 \end{array}$$

Binary carry lookahead adder

Re-expressing the carry equations

$$C_1 = G_0 + P_0C_0$$

$$C_2 = G_1 + P_1C_1 = G_1 + P_1G_0 + P_1P_0C_0$$

$$C_3 = G_2 + P_2C_2 = G_2 + P_2G_1 + P_2P_1G_0 + P_2P_1P_0C_0$$

$$C_4 = G_3 + P_3C_3 = G_3 + P_3G_2 + P_3P_2G_1 + P_3P_2P_1G_0 + P_3P_2P_1P_0C_0$$

All carry bits can be calculated straight from the inputs

3 gate-delays regardless of number of bits, at the cost of more complex logic

In practice, gates have limited number of inputs

For larger adders, can cascade 4-bit CLAs and connect to a group lookahead carry unit

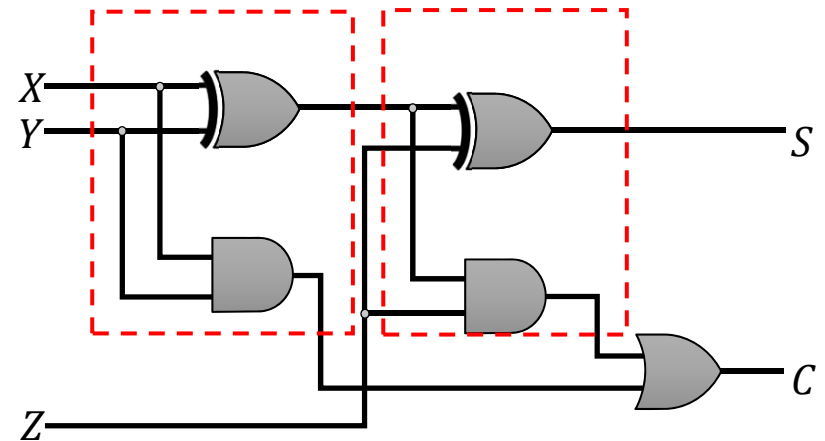


Dataflow Description of 4-Bit Adder

// 4-bit ripple carry adder: dataflow Verilog description
// based on logic diagram earlier in slides

```
module half_adder(X, Y, C, S);  
    input  X, Y;  
    output S, C;  
  
    assign S = X ^ Y;  
    assign C = X & Y;  
  
endmodule
```

```
module full_adder(X, Y, Z, C, S);  
    input  X, Y, Z;  
    output S, C;  
    wire HS, HC, TC;  
  
    half_adder HA1(X, Y, HC, HS);  
    half_adder HA1(HS, Z, TC, S);  
  
    assign C = TC | HC;  
  
endmodule
```



Dataflow Description of 4-Bit Adder

// continued from previous slide

```
module adder_4bit(A, B, C0, S, C4);  
  
    input  [3:0] A, B;  
    input  C0;  
    output [3:0] S;  
    output C4;  
  
    wire [3:1] C;  
  
    full_adder Bit0(A[0], B[0], C0, C[1], S[0]);  
    full_adder Bit1(A[1], B[1], C[1], C[2], S[1]);  
    full_adder Bit2(A[2], B[2], C[2], C[3], S[2]);  
    full_adder Bit3(A[3], B[3], C[3], C4, S[3]);  
  
endmodule
```

Behavioural description of 4-bit adder

// 4-bit full adder: behavioural Verilog description

```
module adder_4_bit(A, B, C0, S, C4);
```

```
    input [3:0] A, B;
```

```
    input C0;
```

```
    output [3:0] S;
```

```
    output C4;
```

```
    assign {C4, S} = A + B + C0;
```

```
endmodule
```

Unsigned binary subtraction

Subtraction involves comparing the subtrahend with the minuend and subtracting the smaller from the larger

If the minuend is larger than the subtrahend, then the result is a positive number; otherwise, negative

Subtraction with comparison is inefficient as comparison operation results in costly circuitry

As an alternative, we can simply subtract the subtrahend from the minuend and correct the result if negative

Unsigned binary subtraction

Example

$$\begin{array}{r} 11110 \\ -10011 \\ \hline \end{array} \qquad \begin{array}{r} 10011 \\ -11110 \\ \hline \end{array}$$

If no borrow occurs into the most significant position, then we know that the subtrahend is not larger than the minuend and the result is positive and correct

If a borrow does occur into the most significant position, then we know the subtrahend is larger than the minuend. The result must then be negative, and so needs correction

Unsigned binary subtraction

The result when a borrow occurs is

$$M - N + 2^n$$

Instead the result we desire is $N - M$ in magnitude. The correct result can be obtained by

$$2^n - (M - N + 2^n) = N - M$$

In the previous example, the correct magnitude is
 $100000 - 10101 = 01011$

Unsigned binary subtraction

In general, the subtraction of two n -digit numbers, $M - N$, in base 2 can be done as follows:

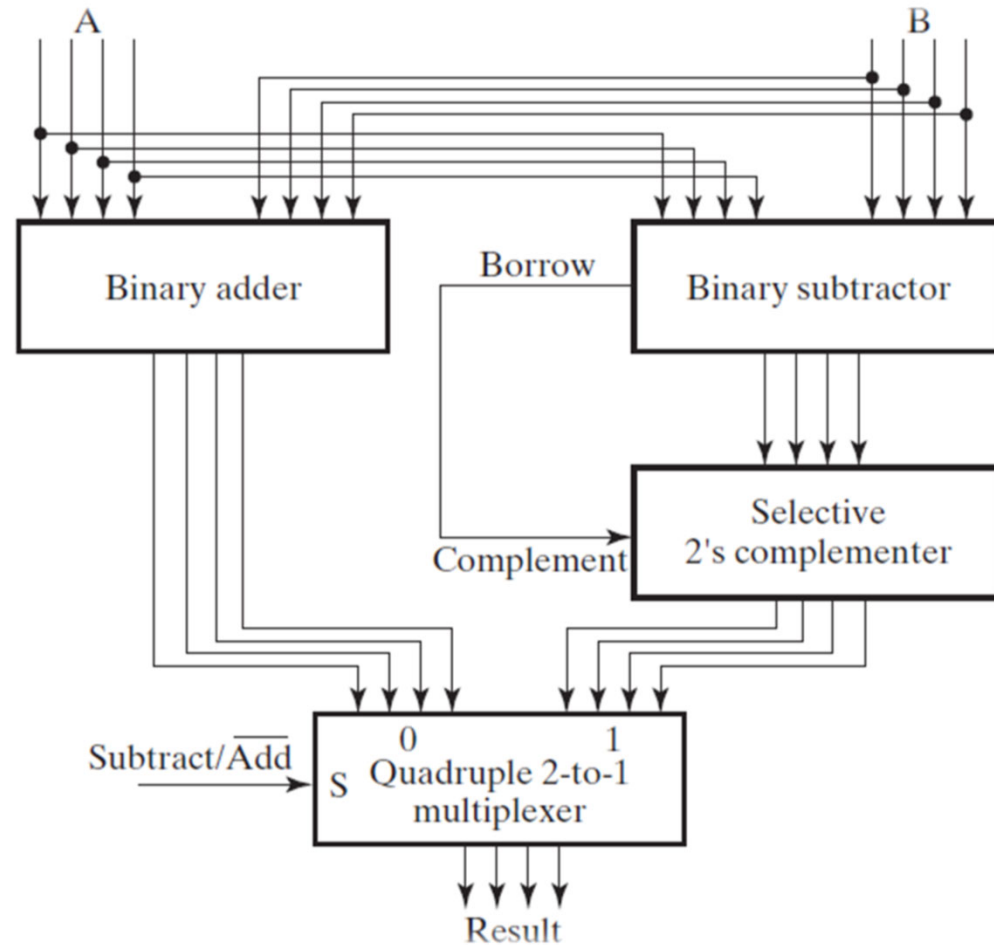
1. Subtract the subtrahend N from the minuend M
2. If no end borrow occurs, then $M \geq N$, and the result is nonnegative and correct
3. If an end borrow occurs, then $N > M$, and the difference, $M - N + 2^n$, is subtracted from 2^n , and a minus sign is appended to the result

The subtraction $2^n - N$ is called taking the 2's complement of N

To do both unsigned addition and unsigned subtraction requires a complex circuit (next slide)



Unsigned binary subtraction



Need shared simpler logic for addition and subtraction

Complements are used to achieve that

Complements

For any radix r , the *Diminished Radix Complement* - called $(r - 1)$'s complement for radix r - is defined as $(r^n - 1) - N$, for some number N

The *Radix Complement* - called r 's complement for radix r is defined as $r^n - N$, for some number N

Example

$$r = 2, N = 0111\ 0011, n = 8$$

$$r^n - 1 = 100000000 - 1 = 11111111$$

The 1 's complement of N can be easily obtained by inverting all the bits in N

$$\begin{array}{r} 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \\ -0\ 1\ 1\ 1\ 0\ 0\ 1\ 1 \\ \hline \end{array}$$

Complements

Example

$$r = 2, N = 0111\ 0011, n = 8$$

$$r^n = 100000000$$

So the **2's complement** of N

$$\begin{array}{r} 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0 \\ -\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1 \\ \hline \end{array}$$

Note the result is the 1's complement plus 1, a fact that can be used in designing hardware

By using complements, it is possible to simplify hardware by sharing adder and subtractor logic

Problem

Obtain the 1's and 2's complement of the following numbers

10011100 10101000 11001101 01101010

Complements

The algorithm for subtracting two unsigned binary numbers using addition

1. Add the 2's complement of the subtrahend N to the minuend M . This performs $M + (2^n - N) = M - N + 2^n$
2. If $M \geq N$, the sum produces an end carry, 2^n . Discard the end carry, leaving result $M - N$
3. If $M < N$, the sum does not produce an end carry, since it is equal to $2^n - (N - M)$, the 2's complement of $N - M$. Perform a correction, taking the 2's complement of the sum and placing a minus sign in front to obtain the result $-(N - M)$

Complements

Example: Given $X = 1010100$ and $Y = 1000011$, perform the subtraction $X - Y$ and $Y - X$

$$\begin{array}{r}
 \quad \quad \quad \quad \quad \quad \quad \\
 X = \\
 2\text{'s complement of } Y = +0 \\
 \hline
 \text{Sum} =
 \end{array}$$

$$\begin{array}{r}
 \quad \quad \quad \quad \quad \quad \quad \\
 Y = \\
 2\text{'s complement of } X \\
 = + \\
 \hline
 \text{Sum} =
 \end{array}$$

$$Y - X = -(2\text{'s complement of } 1101111) =$$

Signed binary numbers

So far we dealt with unsigned binary numbers

Positive numbers and zero are represented in the “usual” way

Negative numbers are identified by adding a minus sign on paper, but no digital implementation

Need to represent negative numbers in hardware

This is done by defining the most significant bit to be the *sign* bit

Use *0* for positive numbers and *1* for negative numbers

Possible ways to represent integers:
Signed-magnitude or *Signed-complement*

Signed-magnitude

The MSB stands for the sign (0 for +, 1 for -)

The remaining bits are of positive magnitude

8-bit example

$$00011001_2 = +25_{10}$$

$$10100101_2 = -37_{10}$$

Arithmetic using signed-magnitude is complex

It requires checking of sign bits and choosing between addition or subtraction

The carry-out or borrow determines whether a correction step should be applied to the result

Signed-complement

Negative numbers are the complements of the respective positive ones

Two possibilities:

1. **Signed 1's complement** - uses 1s complement arithmetic
2. **Signed 2's complement** - uses 2s complement arithmetic

Both will make the MSB correspond to the sign

With 2's complement arithmetic we can use the same hardware as for unsigned numbers alone

Signed-complement

Example: 3-bit numbers

Number	Sign-mag.	1's Comp.	2's Comp.
+3	011	011	011
+2	010	010	010
+1	001	001	001
+0	000	000	000
-0	100	111	-
-1	101	110	111
-2	110	101	110
-3	111	100	101
-4	-	-	100

2's complement arithmetic

2's complement arithmetic does not require special hardware - can use simple adders

The addition of two signed binary numbers with negative numbers represented in signed 2s complement form is obtained from the addition of the two numbers

A carry-out of the MSB bit position is discarded

8 bit example:

$$\begin{array}{r} +6 \quad 00000110 \\ +13 \quad +00001101 \\ \hline +19 \end{array}$$

$$\begin{array}{r} +6 \quad 00000110 \\ -13 \quad +11110011 \\ \hline -7 \end{array}$$

$$\begin{array}{r} -6 \quad 11111010 \\ +13 \quad +00001101 \\ \hline +7 \end{array}$$

$$\begin{array}{r} -6 \quad 11111010 \\ -13 \quad +11110011 \\ \hline -19 \end{array}$$



2's complement arithmetic

To convert negative binary numbers to decimal, use one of two methods:

1. Take the 2's complement of the negative number to obtain the corresponding positive one, and add a minus sign at the front
2. Convert as usual but subtract the MSBs value instead of adding

Example: for the 8-bit number 1110 1101:

1. 2's complement:
2. $1110\ 1101_2$

2's complement subtraction

Subtraction with 2's complement is simple

Take the 2s complement of the subtrahend and add it to the minuend

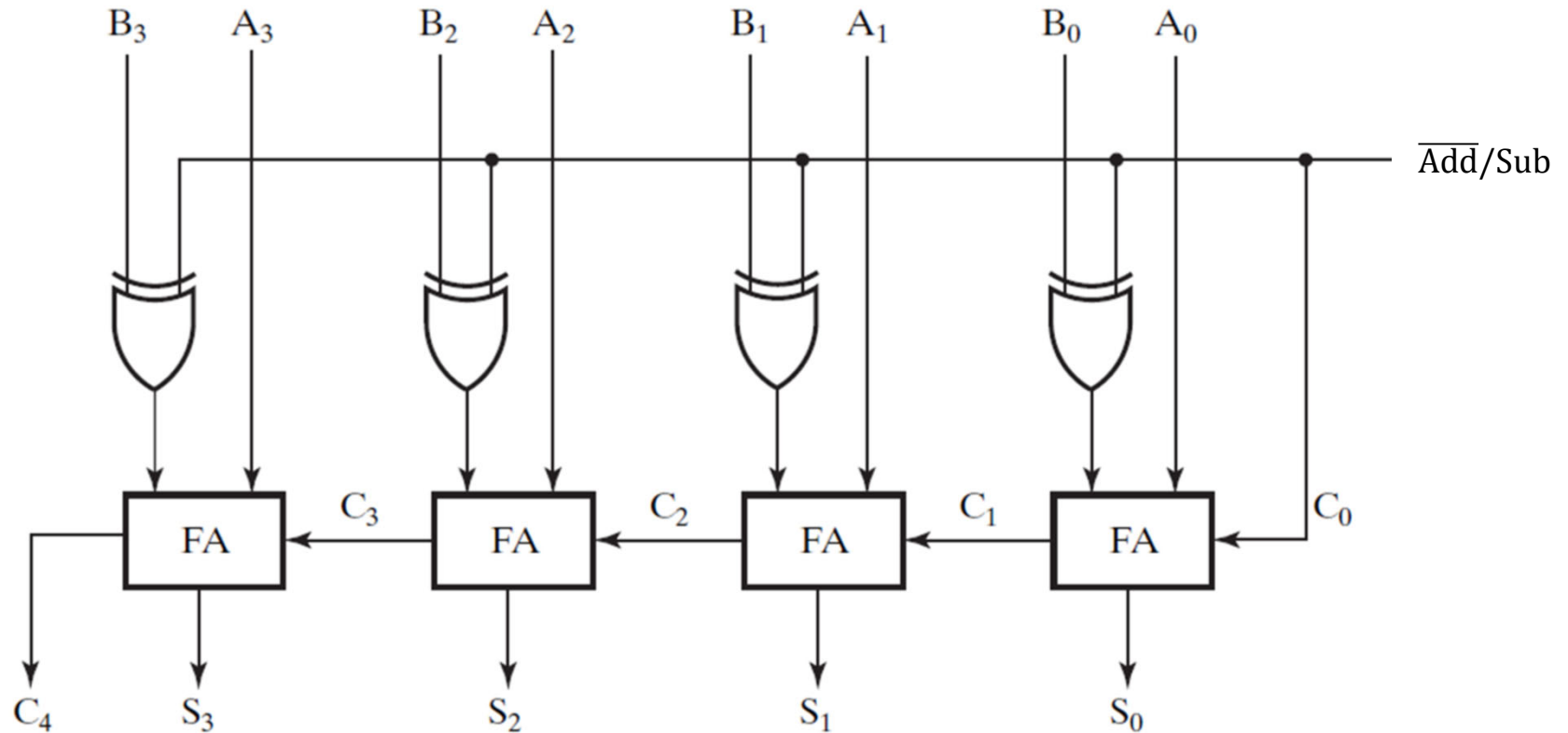
A carry-out of the sign bit position is discarded

Works because: $A - B = A + (-B)$

8-bit example:

$$\begin{array}{r} -6 \quad 11111010 \\ - -13 \quad -11110011 \\ \hline +7 \end{array} \quad \longrightarrow \quad \begin{array}{r} 11111010 \\ +00001101 \\ \hline \end{array}$$

2's complement adder/subtractor



$$\begin{aligned} S &= A + B & \overline{\text{Add/Sub}} &= 0 \\ S &= A - B & \overline{\text{Add/Sub}} &= 1 \end{aligned}$$

2's complement adder/subtractor

Signal $\overline{\text{Add/Sub}}$ selects between addition or subtraction

If $\overline{\text{Add/Sub}} = 0$, vector B propagate through the XOR gates and the carry-in is 0 - computes $A + B$

If $\overline{\text{Add/Sub}} = 1$, vector B is complemented and the carry-in is 1 to obtain the 2s complement of B - computes $A - B$

Sign extension

To represent an n -bit number with larger number of bits ($n + m$):

For *unsigned* numbers add m 0's at the front

For *signed* numbers, extend the MSB at the front of the number

4 to 8 bits signed example:

$$+7 = 0111 = 0000\ 0111$$

$$-7 = 1001 = 1111\ 1001$$

Problem

The following signed binary number are in 2's complement form. Perform the following operations

1. $00000111 + 11110100$

2. $0110111 - 1010111$

Overflow

Overflow occurs if $(n + 1)$ bits are required to contain the result from an n -bit addition or subtraction

Example: 8-bits can represent values between -128 to +127.
Calculate $70 + 80$:

$$\begin{array}{r} 0\ 1\ 0\ 0\ 0\ 1\ 1\ 0 \\ + 0\ 1\ 0\ 1\ 0\ 0\ 0\ 0 \\ \hline 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0 \end{array}$$

Negative number! (MSB = 1)

For *unsigned* addition:

$C = 0 \rightarrow$ No Overflow

$C = 1 \rightarrow$ Overflow

For *signed* addition (and subtraction):

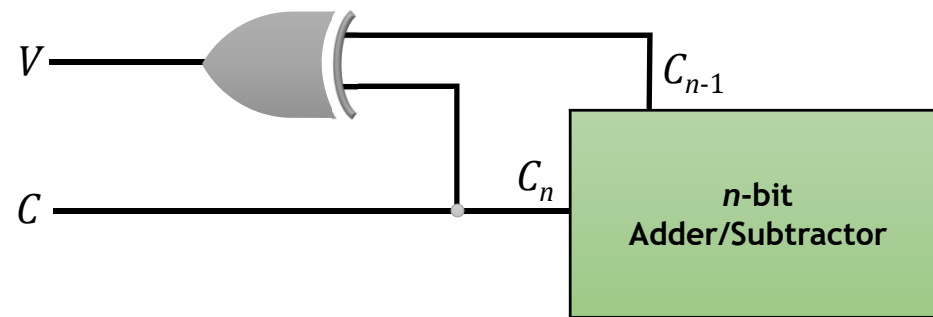
$V = 0 \rightarrow$ No Overflow

$V = 1 \rightarrow$ Overflow

Overflow

Overflow occurs when the carry-in to the MSB is not equal to the carry-out from the MSB

A simple circuit to detect overflow



N, Z, C, V - Status flags

Status signals that provide information from the arithmetic unit in microprocessors

N (*Negative*): 1 if the sum is negative; 0 if positive

Z (*Zero*): 1 if the sum is zero; 0 if non-zero

C (*Carry*): 1 if there is a carry-out; 0 if no carry

V (*oVerflow*): 1 if signed overflow detected; 0 for no-overflow

Can be used to compare numbers

N, Z, C, V - Status flags

These can be implemented by:

$$N = S_{n-1} \quad (\text{value at MSB of sum})$$

$$Z = \overline{S_{n-1} + S_{n-2} + \dots + S_0} \quad (\text{detect if all sum bits are 0})$$

$$C = C_n \quad (\text{carry out from MSB})$$

$$V = C_n \oplus C_{n-1} \quad (\text{MSB carry-in} \neq \text{carry-out})$$

$$\begin{array}{r} 0 \ 1 \ 0 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0 \\ \ 0 \ 1 \ 0 \ 1 \ 0 \ 0 \ 1 \ 1 \\ + 0 \ 1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \\ \hline 1 \ 0 \ 1 \ 1 \ 1 \ 0 \ 0 \ 1 \end{array}$$

$$N = 1; Z = 0; C = 0; V = 1$$

N, Z, C, V - Status flag

// status flags

```
module add32_status_flags(a, b, s, c, z, n, v);  
    input reg signed [31:0] a,b;  
    output reg signed [31:0] s;  
    output z,n,v,c;  
    assign {c,s} = a + b;  
    assign z = ~|s;  
    assign n = s[31];  
    assign v = a[31]^b[31]^s[31]^c;  
endmodule
```

BCD addition

BCD addition is complicated by the fact that the sum can exceed 9

$$\begin{array}{r}
 \begin{array}{l} X \\ + Y \\ \hline Z \end{array} \quad \begin{array}{l} 0111 \\ + 0101 \\ \hline 1100 \\ + 0110 \\ \hline 10010 \end{array} \quad \begin{array}{l} 7 \\ + 5 \\ \hline 12 \end{array} \\
 \text{carry} \rightarrow \underbrace{10010}_{S=2}
 \end{array}$$

$$Z = X + Y$$

If $Z \leq 9$, then $S = Z$ & $C = 0$

If $Z \geq 9$, then $S = Z + 6$ & $C = 1$

$$\begin{array}{r}
 \begin{array}{l} X \\ + Y \\ \hline Z \end{array} \quad \begin{array}{l} 1000 \\ + 1001 \\ \hline 10001 \\ + 0110 \\ \hline 10111 \end{array} \quad \begin{array}{l} 8 \\ + 9 \\ \hline 17 \end{array} \\
 \text{carry} \rightarrow \underbrace{10111}_{S=7}
 \end{array}$$

$$Z = X + Y$$

If $Z \geq 16$, then $S = Z + 6$ & $C = 1$

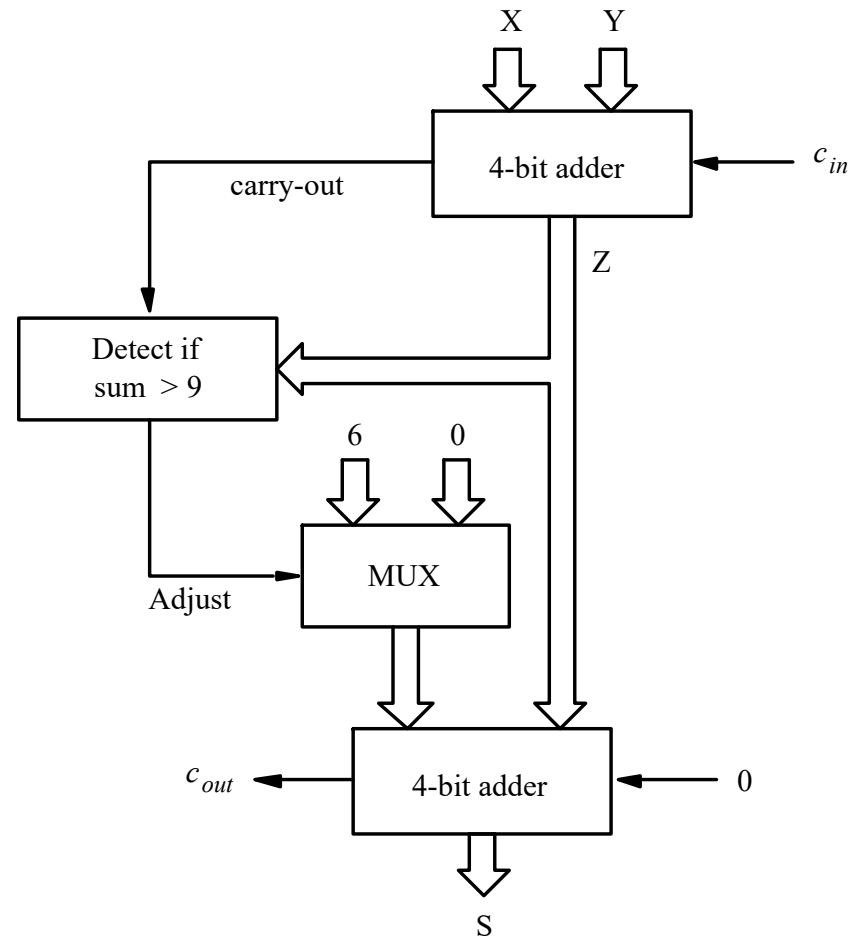


UNSW
SYDNEY

BCD addition

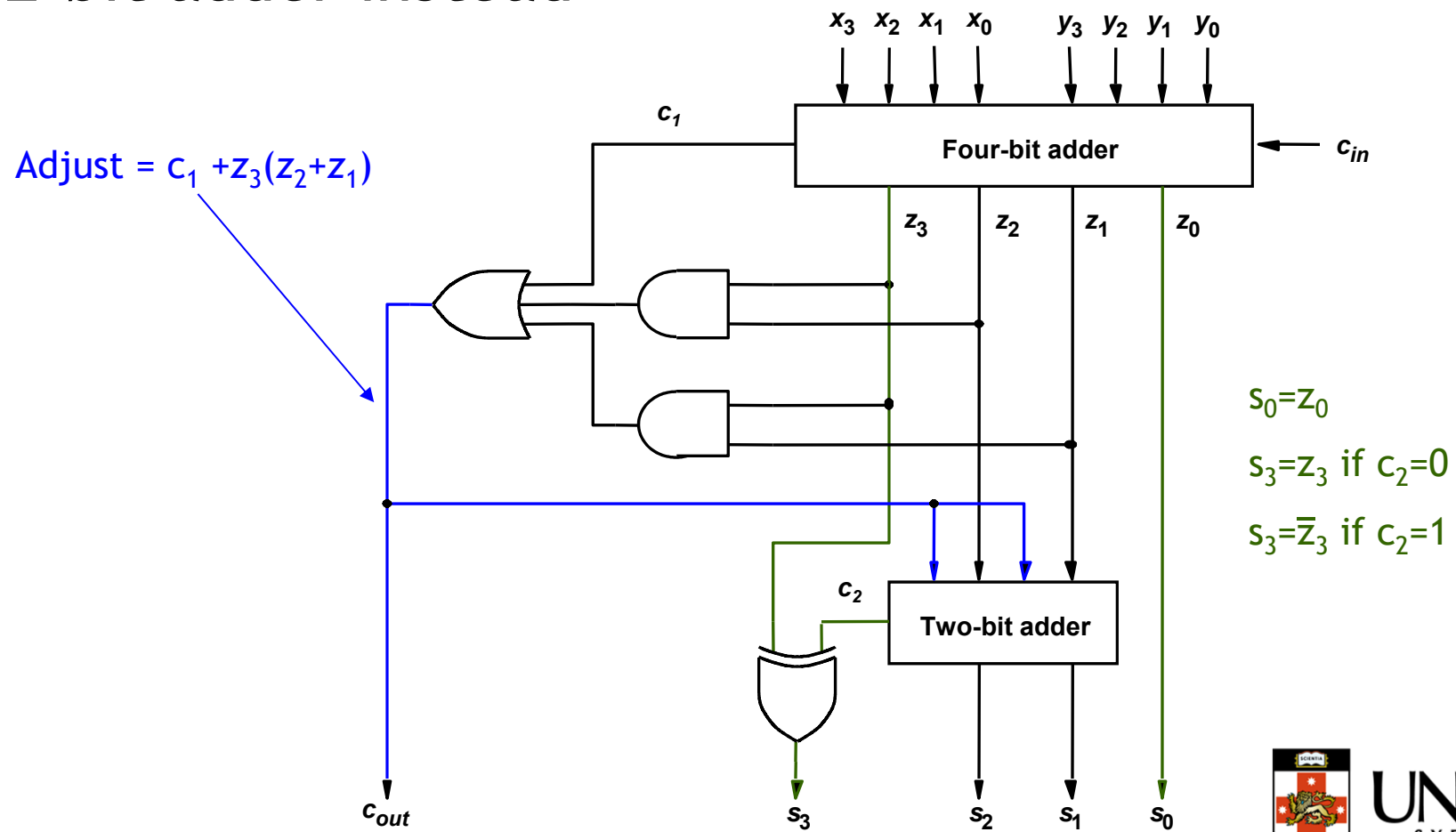
As the correction for when the sum is greater than 9 or 16 is the same, only need to check whether $\text{sum} > 9$

The signal Adjust controls whether the MUX provides correction or not



BCD addition

As 6 is 0110, the correction can be done using a 2-bit adder instead



Binary multiplication

The binary digit multiplication table is trivial

$a \times b$	$b = 0$	$b = 1$
$a = 0$	0	0
$a = 1$	0	1

This is simply the Boolean AND function

Larger binary products are formed just like larger products are formed in base 10

Binary multiplication

In decimal multiplication, partial products are computed and then summed up

The partial product summation for n digits, requires adding up to n digits (with carries)

Also an $n \times m$ digits multiply generates up to an $m + n$ digit result

$$\begin{array}{r} \\ \\ \\ \\ + \\ \hline 3 \end{array}$$

Binary multiplication

Multiplying a number by 2^k , shifts it to the left by k bit positions

Dividing a number by 2^k , shifts it to the right by k bit positions

For unsigned numbers 0s are added to the left of the MSB

In signed numbers, as it is necessary to preserve the sign, bits are shifted to the right and the value of the sign bit filled from the left

A = 011000

B = 101000

Binary multiplication

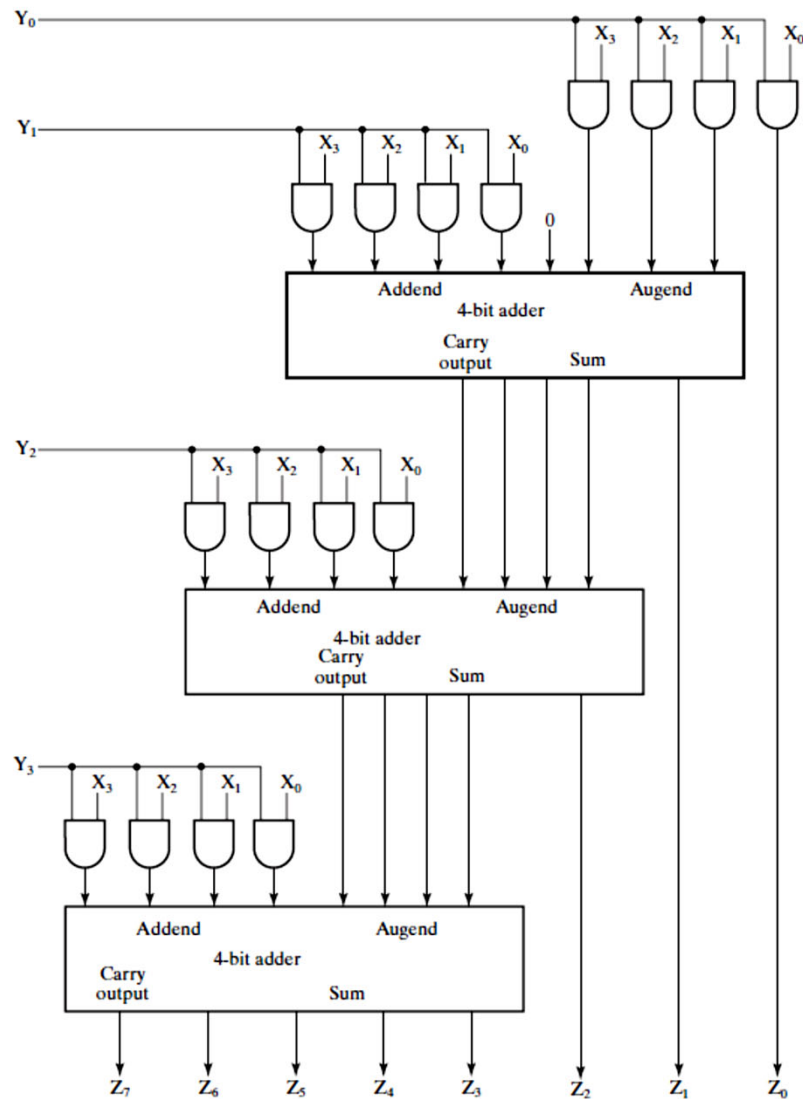
Multiplicand X	(14)	1 1 1 0
Multiplier Y	(11)	x 1 0 1 1
		<hr/>
		1 1 1 0
		1 1 1 0
		0 0 0 0
		1 1 1 0
		<hr/>
Product Z	(154)	1 0 0 1 1 0 1 0

Partial products are

either all zero (if the multiplier digit is zero) or
the same as the multiplicand (if the multiplier digit is one)

Note: No carries are added in partial product formation

Binary multiplication



Partial products are formed using an $n \times m$ array of AND gates

The partial products are then summed using adder trees

Will need $m - 1$ adders of width n bits

Binary multiplication

In hardware, a sequential approach is adopted with adders used to compute partial products

Multiplicand M(11)	1 1 1 0
Multiplier Q (14)	X 1 0 1 1
	<hr/>
Partial product 0	1 1 1 0
	+ 1 1 1 0
	<hr/>
Partial product 1	1 0 1 0 1
	+ 0 0 0 0
	<hr/>
Partial product 2	0 1 0 1 0
	+ 1 1 1 0
	<hr/>
Product P (154)	1 0 0 1 1 0 1 0

Diagram illustrating binary multiplication using adders. The multiplicand M(11) is 1110 and the multiplier Q(14) is 1011. The partial products are: Partial product 0 (1110), Partial product 1 (10101), and Partial product 2 (01010). The final product P(154) is 10011010. Blue arrows indicate the sequential addition of partial products to the running total.

Faster to use multiple adders instead of a single adder

Binary multiplication

Sequential multiplication circuits are slow

Can be made faster by performing addition in array form

$M = m_3 m_2 m_1 m_0$ ← 4x4 multiplication
 $Q = q_3 q_2 q_1 q_0$ ←
 $PP0 = m_3 q_0 \ m_2 q_0 \ m_1 q_0 \ m_0 q_0$

$$\begin{array}{rcccl}
PP0: & 0 & pp0_3 & pp0_2 & pp0_1 & pp0_0 \\
& + & m_3q_1 & m_2q_1 & m_1q_1 & m_0q_1 & 0 \\
\hline
PP1: & pp1_4 & pp1_3 & pp1_2 & pp1_1 & pp1_0
\end{array}$$

Binary multiplication

