

Unit 2

Integer Operations

(Arithmetic, Overflow, Bitwise Logic, Shifting)

Skills & Outcomes

- You should know and be able to apply the following skills with confidence
 - Perform addition & subtraction in unsigned & 2's complement system
 - Determine if overflow has occurred
 - Perform bitwise operations on numbers
 - Perform logic and arithmetic shifts and understand how they can be used for multiplication/division
 - Understand arithmetic in binary and hex

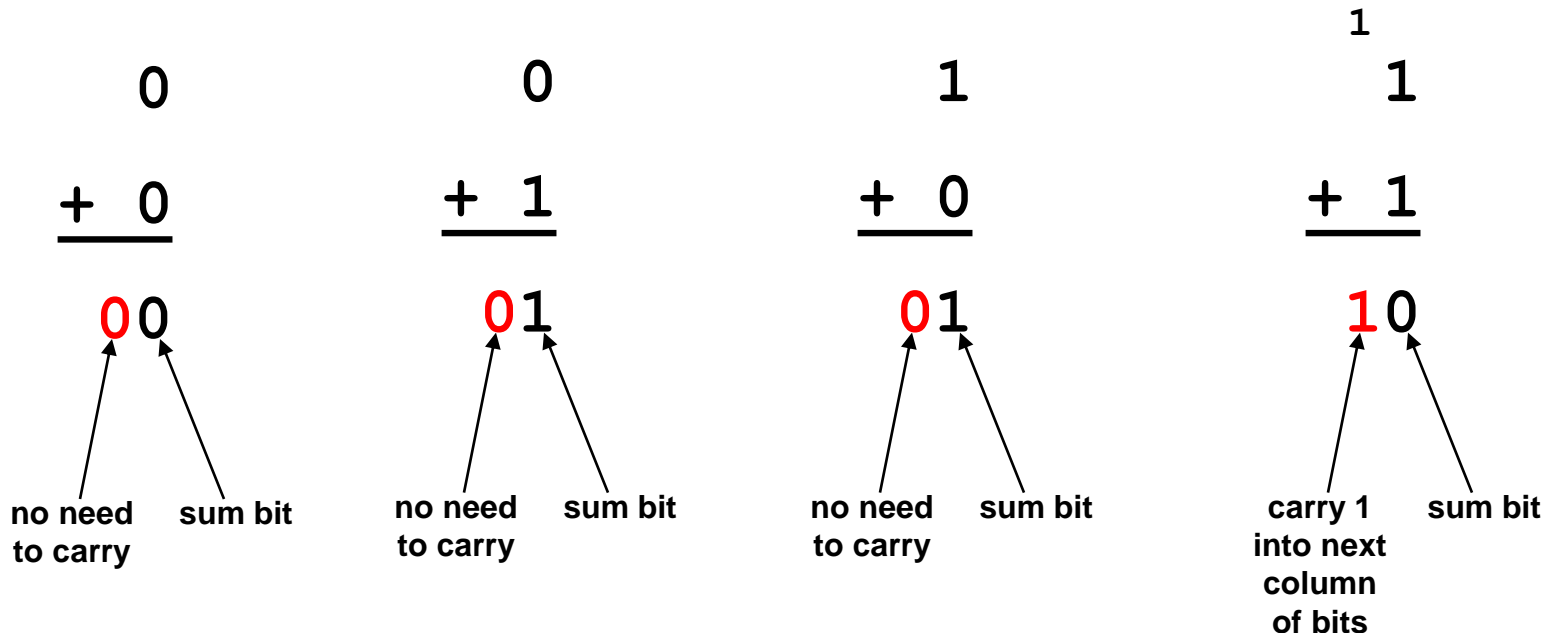
UNSIGNED BINARY ARITHMETIC

Binary Arithmetic

- Can perform all arithmetic operations (+, -, *, ÷) on binary numbers
- Can use same methods as in decimal
 - Still use carries and borrows, etc.
 - Only now we carry when sum is 2 or more rather than 10 or more (decimal)
 - We borrow 2's not 10's from other columns
- Easiest method is to add bits in your head in decimal ($1+1 = 2$) then convert the answer to binary ($2_{10} = 10_2$)

Binary Addition

- In decimal addition we carry when the sum is 10 or more
- In binary addition we carry when the sum is 2 or more
- Add bits in binary to produce a sum bit and a carry bit



Binary Addition & Subtraction

$$\begin{array}{r}
 1\ 1\ 1 \\
 0\ 1\ 1\ 1\ (7) \\
 +\ 0\ 0\ 1\ 1\ (3) \\
 \hline
 1\ 0\ 1\ 0\ (10) \\
 \begin{array}{cccc}
 8 & 4 & 2 & 1
 \end{array}
 \end{array}$$

$$\begin{array}{r}
 0\ 0 \\
 \cancel{1}\ 10\ \cancel{1}\ 10\ (10) \\
 -\ 0\ 1\ 0\ 1\ (5) \\
 \hline
 0\ 1\ 0\ 1\ (5) \\
 \begin{array}{cccc}
 8 & 4 & 2 & 1
 \end{array}
 \end{array}$$

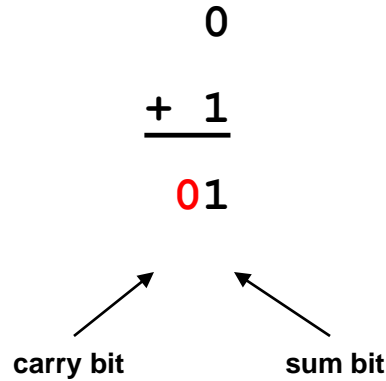
Binary Addition

$$\begin{array}{r} 110 \\ 0110 \text{ (6)} \\ 8 4 2 1 \\ + 0111 \text{ (7)} \\ \hline 1101 \text{ (13)} \end{array}$$

Binary Addition

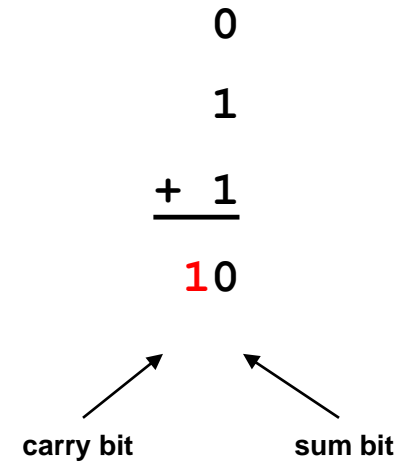
1

$$\begin{array}{r}
 0 \\
 0110 \text{ (6)} \\
 + 0111 \text{ (7)} \\
 \hline
 1101 \text{ (13)}
 \end{array}$$



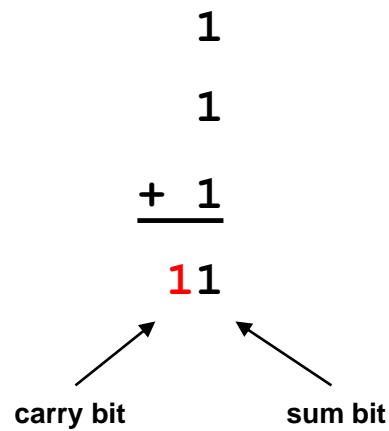
2

$$\begin{array}{r}
 10 \\
 0110 \text{ (6)} \\
 + 0111 \text{ (7)} \\
 \hline
 1101 \text{ (13)}
 \end{array}$$



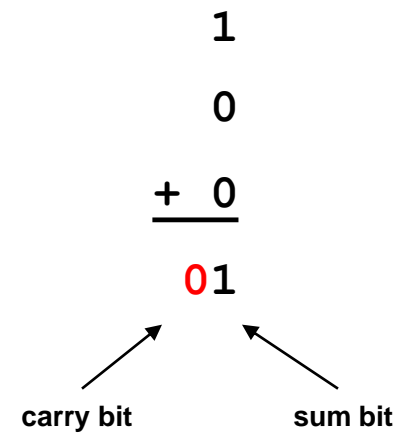
3

$$\begin{array}{r}
 110 \\
 0110 \text{ (6)} \\
 + 0111 \text{ (7)} \\
 \hline
 1101 \text{ (13)}
 \end{array}$$



4

$$\begin{array}{r}
 110 \\
 0110 \text{ (6)} \\
 + 0111 \text{ (7)} \\
 \hline
 1101 \text{ (13)}
 \end{array}$$



Hexadecimal Arithmetic

- Same style of operations
 - Carry when sum is 16 or more, etc.

$$\begin{array}{r} 1 \ 1 \\ 4 \ D_{16} \end{array}$$

$$\begin{array}{r} + \ B \ 5_{16} \\ \hline \end{array}$$

$$\begin{array}{r} 1 \ 0 \ 2_{16} \end{array}$$

$$13+5 = 18_{10} = \underline{1} \underline{2}_{16}$$

16 1

$$1+4+11 = 16_{10} = \underline{1} \underline{0}_{16}$$

16 1

Binary Multiplication

- Like decimal multiplication, find each partial product and shift them, then sum them up
- Multiplying **two n -bit numbers** yields at most a **$2*n$ -bit product**

				0	1	1	0	(6)					
					*	0	1	0	1	(5)			
						<hr/>							
						0	1	1	0	} Partial Products			
						0	0	0	0				
						0	1	1	0				
						0	0	0	0				
						<hr/>							
						0	0	1	1	1	1	0	← Sum of the partial products

Binary Division

- Use the same long division techniques as in decimal
- Dividing **two n -bit numbers** may yield an **n -bit quotient** and **n -bit remainder**

$$\begin{array}{r}
 (2)_{10} \quad 10 \quad \overline{) \begin{array}{cccc}
 0 & 1 & 0 & 1 \\
 \hline
 1 & 0 & 1 & 1 \\
 \hline
 -1 & 0 & & \\
 \hline
 0 & 1 & & \\
 -0 & 0 & & \\
 \hline
 & 1 & 1 & \\
 & -1 & 0 & \\
 \hline
 & & 0 & 1 \\
 \hline
 & & & 1 \\
 \hline
 & & & 0 \\
 \hline
 & & & 1
 \end{array} \\
 \end{array}$$

$(5 \text{ r.} 1)_{10}$
 $(11)_{10}$

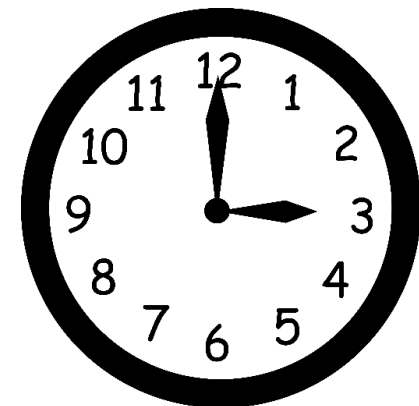
$0 \ 1 \ 0 \ 1 \ \text{r.} \ 1 \ (5 \ \text{r.} \ 1)_{10}$

"Taking the 2's complement"

SUBTRACTION THE EASY WAY

Modulo Arithmetic

- The primary difference between how humans and computers perform arithmetic is the finite precision of computers
 - As humans we can use more digits (precision) as needed
 - Computers can only use a finite set of bits
 - Much like the odometer on your car once you go too many miles the values will wrap from 999999 to 000000
 - Essentially all computer arithmetic is modulo arithmetic
 - If we have a width of w bits, then all operations are module 2^w
- This leads to alternate approaches to arithmetic
 - Example: Consider how you could change the clock time from 5 p.m. to 3 p.m. if you can't subtract hours



Taking the Negative

CS:APP 2.3.3

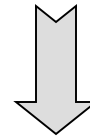
- **Question:** Given a number in 2's complement how do we find its negative (i.e. $-1 * X$)
- **Answer:** By "taking the 2's complement"
 - $0110 = +6 \Rightarrow -6 = 1010$
 - Operation defined as:
 1. Flip/invert/not all the bits (1's complement)
 2. Add 1 and drop any carry (i.e. finish with the same # of bits as we start with)
 - See next slides for example

Taking the 2's Complement

- Invert (flip) each bit (take the 1's complement)
 - 1's become 0's
 - 0's become 1's
- Add 1 (drop final carry-out, if any)

-32 16 8 4 2 1
010011

Original number = +19



101100

Bit flip is called the 1's complement of a number

+ **1**

Resulting number = -19

101101

Important: Taking the 2's complement is equivalent to taking the negative (negating)

Taking the 2's Complement

①

-32 16 8 4 2 1	Original number = -22
101010	

010101	Take the 2's complement yields the negative of a number
+ 1	
010110	

Resulting number = +22

101001	Taking the 2's complement again yields the original number (the operation is symmetric)
+ 1	
101010	

Back to original = -22

②

0000	Original # = 0
-------------	----------------

1111	Take the 2's complement
+ 1	
0000	

2's comp. of 0 is 0

③

1000	Original # = -8
-------------	-----------------

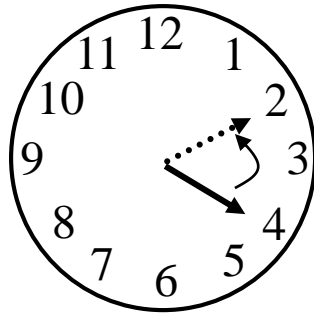
0111	Take the 2's complement
+ 1	
1000	

Negative of -8 is -8
(i.e. no positive
equivalent, but this is
not a huge problem)

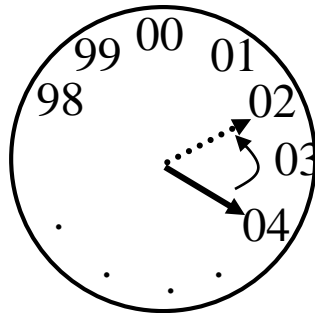
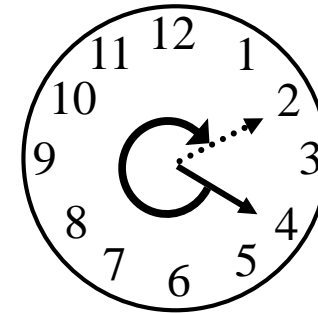
The same algorithms regardless of unsigned or signed

ADDITION AND SUBTRACTION

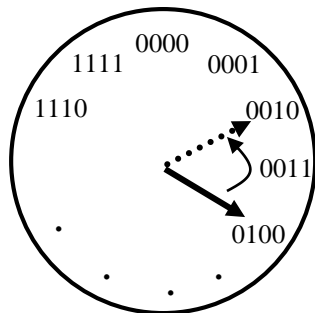
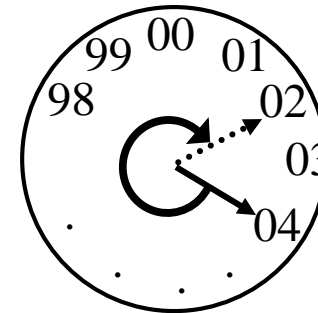
Radix Complement



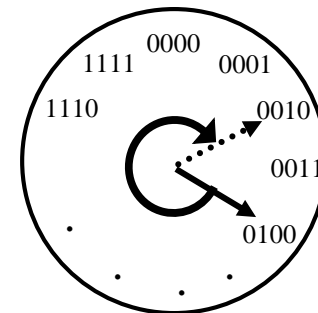
Clock Analogy
 $4 - 2 = 4 + 10$



10's complement
 $04 - 02 = 04 + 98$



2's complement
 $0100 - 0010 = 0100 + 1110$



When using **modulo arithmetic**, **subtraction** can always be converted to **addition**.

2's Complement Addition/Subtraction

CS:APP 2.3.1
CS:APP 2.3.2

- Addition
 - Sign of the numbers do not matter
 - Add column by column
 - Drop any final carry-out
 - The secret to modulo arithmetic
- Subtraction
 - Any subtraction $(A-B)$ can be converted to addition $(A + -B)$ by taking the 2's complement of B
 - $(A-B)$ becomes $(A + \sim B + 1)$
 - Drop any carry-out
 - The secret to modulo arithmetic

2's Complement Addition

- No matter the sign of the operands just add as normal
- Drop any extra carry out

$$\begin{array}{r}
 0000 \\
 0011 \text{ (3)} \\
 + 0010 \text{ (2)} \\
 \hline
 0101 \text{ (5)}
 \end{array}$$

$$\begin{array}{r}
 0000 \\
 1101 \text{ (-3)} \\
 + 0010 \text{ (2)} \\
 \hline
 1111 \text{ (-1)}
 \end{array}$$

Drop final carry-out →

$$\begin{array}{r}
 ~~1~~110 \\
 0011 \text{ (3)} \\
 + 1110 \text{ (-2)} \\
 \hline
 0001 \text{ (1)}
 \end{array}$$

$$\begin{array}{r}
 ~~1~~100 \\
 1101 \text{ (-3)} \\
 + 1110 \text{ (-2)} \\
 \hline
 1011 \text{ (-5)}
 \end{array}$$

Unsigned and Signed Addition

- Addition process is the same for both unsigned and signed numbers
 - Add columns right to left
- Examples:


1 1	<u>If unsigned</u>	<u>If signed</u>
1001	(9)	(-7)
+ 0011	(3)	(3)
1100	(12)	(-4)

2's Complement Subtraction

- Take the 2's complement of the subtrahend (bottom #) and add to the original minuend (top #)
- Drop any extra carry out

$$\begin{array}{r} 0011 (+3) \\ - 0010 (+2) \\ \hline \end{array}$$

$$\begin{array}{r} 1101 (-3) \\ - 1110 (-2) \\ \hline \end{array}$$

Drop final carry-out 

$$\begin{array}{r} 1111 \\ 0011 \\ 1101 \text{ Bit flip of } +2 \\ + \quad 1 \text{ Add } 1 \\ \hline 0001 \end{array}$$

$$\begin{array}{r} 1 \\ 1101 \\ 0001 \text{ Bit flip of } -2 \\ + \quad 1 \text{ Add } 1 \\ \hline 1111 \end{array}$$

Unsigned and Signed Subtraction

- Subtraction process is the same for both unsigned and signed numbers
 - Convert $A - B$ to $A + \text{Comp. of } B$
 - Drop any final carry out
- Examples:

	<u>If unsigned</u>	<u>If signed</u>					
1100	(12)	(-4)		11 1			
- 0010	(2)	(2)	→	1100		A	
<u> </u>				1101		~B	
				+ 1		Add 1	
				<u> </u>			
				1010		(10)	(-6)
						<u>If unsigned</u>	<u>If signed</u>

Important Note

- Almost all computers use 2's complement because...
- The same addition and subtraction algorithm can be used on unsigned and 2's complement (signed) numbers
- Thus we only need one set of circuitry (HW component) to perform operations on both unsigned and signed numbers

OVERFLOW

Overflow

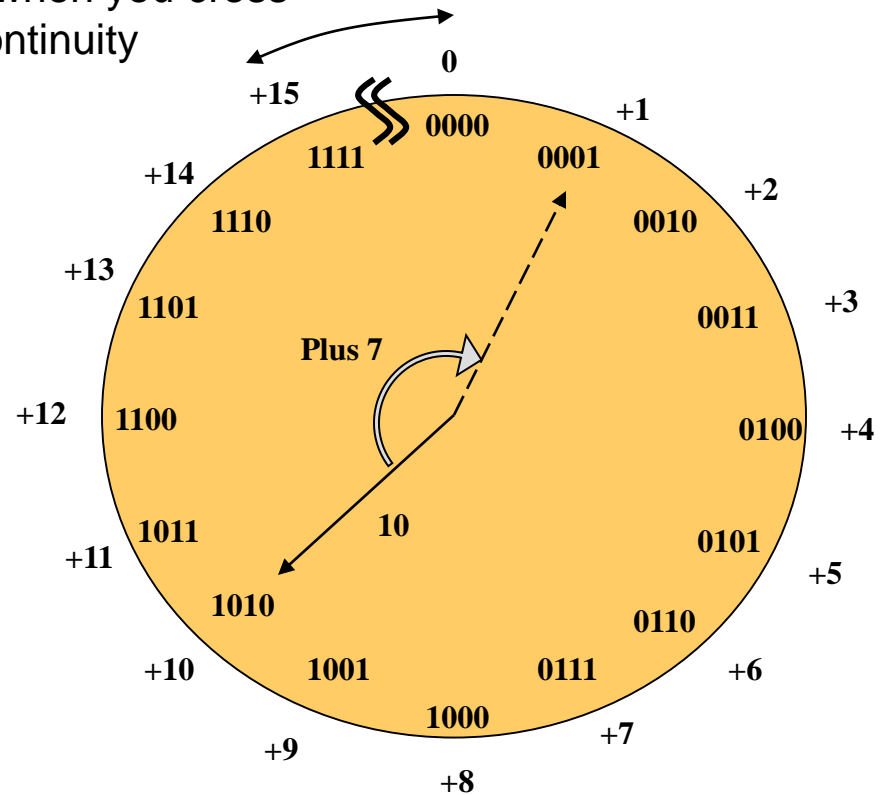
- Overflow occurs when the result of an arithmetic operation is too large to be represented with the given number of bits
- Conditions and tests to determine overflow depend on the system being used
 - Different algorithms for detecting overflow based on unsigned or signed

Unsigned Overflow

Overflow occurs when you cross this discontinuity

$$10 + 7 = 17$$

With 4-bit *unsigned* numbers we can only represent 0 – 15. Thus, we say overflow has occurred.

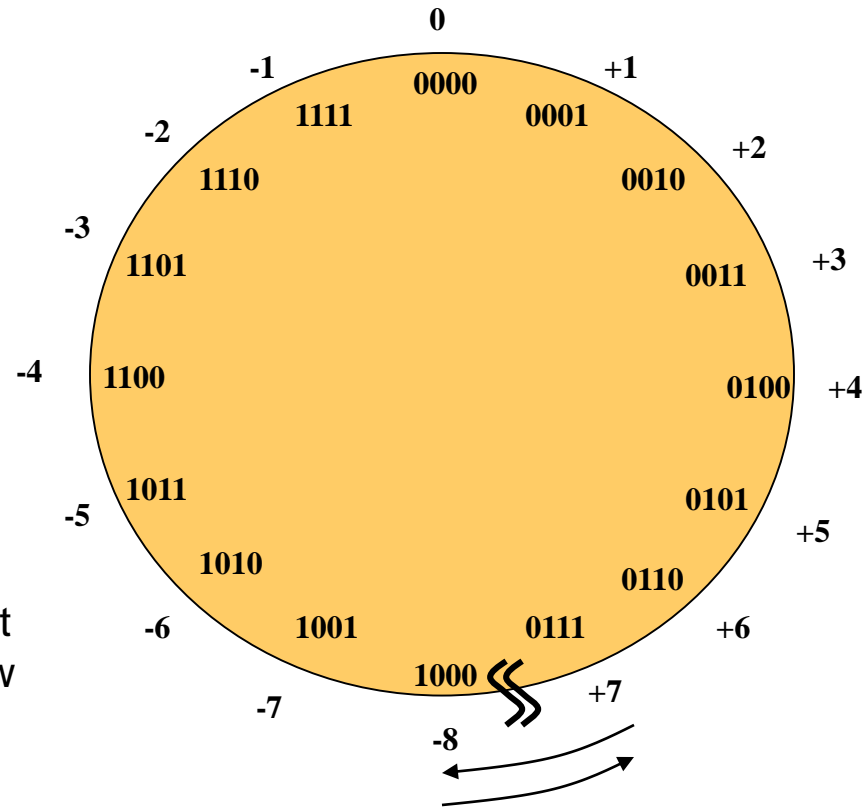


2's Complement Overflow

$$5 + 7 = +12$$

$$-6 + -4 = -10$$

With 4-bit 2's complement numbers we can only represent -8 to +7. Thus, we say overflow has occurred.



Overflow occurs when you cross this discontinuity

Overflow in Addition

- Overflow occurs when the result of the addition cannot be represented with the given number of bits.
- Tests for overflow:
 - Unsigned: if Cout = 1 [result smaller than inputs]
 - Signed: if p+p=n or n+n=p [result has inappropriate sign]

1 1	<u>If unsigned</u>	<u>If signed</u>	0 1	<u>If unsigned</u>	<u>If signed</u>
1101	(13)	(-3)	0110	(6)	(6)
+ 0100	(4)	(4)	+ 0101	(5)	(5)
<hr style="width: 100%; border: 0.5px solid black; margin: 0;"/> 0001	(17)	(+1)	<hr style="width: 100%; border: 0.5px solid black; margin: 0;"/> 1011	(11)	(-5)
	<u>Overflow</u>	<u>No Overflow</u>		<u>No Overflow</u>	<u>Overflow</u>
	Cout = 1	n + p		Cout = 0	p + p = n

Overflow in Subtraction

- Overflow occurs when the result of the subtraction cannot be represented with the given number of bits.
- Tests for overflow:
 - Unsigned: if Cout = 0 [expect negative result]
 - Signed: if p+p=n or n+n=p [result has inappropriate sign]

$\begin{array}{r} 0111 \\ - 1000 \\ \hline \end{array}$	<p style="text-align: center;"><u>If unsigned</u> <u>If signed</u></p> <p style="text-align: center;">(7) (7)</p> <p style="text-align: center;">(8) (-8)</p> <p style="text-align: center;">(-1) (15)</p>		$\begin{array}{r} 0111 \\ 0111 \text{ A} \\ 0111 \text{ 1's comp. of B} \\ + \quad 1 \text{ Add 1} \\ \hline 1111 \end{array}$	<p style="text-align: center;"><u>Desired Results</u></p>	<p style="text-align: center;">(15) (-1)</p>	<p style="text-align: center;"><u>If unsigned Overflow</u></p> <p style="text-align: center;">Cout = 0</p>	<p style="text-align: center;"><u>If signed Overflow</u></p> <p style="text-align: center;">p + p = n</p>
---	--	--	--	---	--	--	---

MULTIPLICATION AND DIVISION

Binary Multiplication

CS:APP 2.3.4

- Multiplying **two n -bit numbers** yields at most a **$2*n$ -bit product**
- Multiplication operations on a modern processor can take **3-5 times** longer than addition operations

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

} Partial Products
← Sum of the partial products

Binary Division

- Dividing **two n -bit numbers** may yield an **n -bit quotient and n -bit remainder**
- Division operations on a modern processor can take **17-41 times** longer than addition operations

$$\begin{array}{r}
 \text{0 1 0 1 r.1} \quad (5 \text{ r.1})_{10} \\
 (2)_{10} \quad 10 \overline{) 1011} \quad (11)_{10} \\
 \underline{-10} \\
 01 \\
 \underline{-00} \\
 11 \\
 \underline{-10} \\
 01
 \end{array}$$

Unsigned Multiplication Review

- Same rules as decimal multiplication
- Multiply each bit of Q by M shifting as you go
- An m-bit * n-bit mult. produces an m+n bit result
- Notice each partial product is a shifted copy of M or 0 (zero)

	1010	M (Multiplicand)
	* 1011	Q (Multiplier)
	1010	
	1010_	PP (Partial
	0000_	Products)
	+ 1010	
	01101110	P (Product)

Signed Multiplication Techniques

- When multiplying signed (2's comp.) numbers, some new issues arise
- Must sign extend partial products (out to 2n bits)

**Without Sign Extension...
Wrong Answer!**

$$\begin{array}{r}
 1001 = -7 \\
 * 0110 = +6 \\
 \hline
 0000 \\
 1001\text{ } \\
 1001\text{ } \\
 + 0000 \\
 \hline
 00110110 = +54
 \end{array}$$

**With Sign Extension...
Correct Answer!**

$$\begin{array}{r}
 1001 = -7 \\
 * 0110 = +6 \\
 \hline
 00000000 \\
 1111001\text{ } \\
 111001\text{ } \\
 + 00000 \\
 \hline
 11010110 = -42
 \end{array}$$

Signed Multiplication Techniques

- Also, must worry about negative multiplier
 - MSB of multiplier has negative weight
 - If MSB=1, multiply by -1 (i.e. take 2's comp. of multiplicand)

**With Sign Extension but w/o consideration of MSB...
Wrong Answer!**

$$\begin{array}{r}
 1100 = -4 \\
 * 1010 = -6 \\
 \hline
 00000000 \\
 1111100\text{_} \\
 000000\text{_} \\
 + 11100\text{_} \\
 \hline
 11011000 = -40
 \end{array}$$

**With Sign Extension and w/ consideration of MSB...
Correct Answer!**

Place Value: -8
Multiply by -1

$$\begin{array}{r}
 1100 = -4 \\
 * \textcircled{1}010 = -6 \\
 \hline
 00000000 \\
 1111100\text{_} \\
 000000\text{_} \\
 + 00100\text{_} \\
 \hline
 00011000 = +24
 \end{array}$$

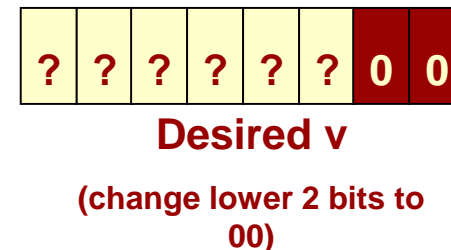
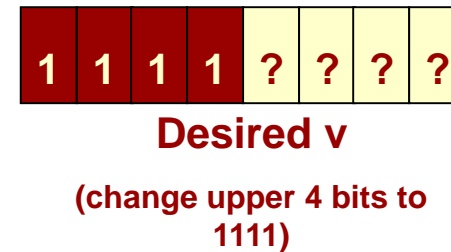
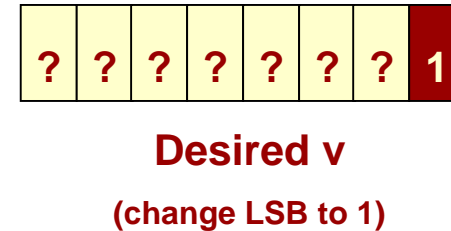
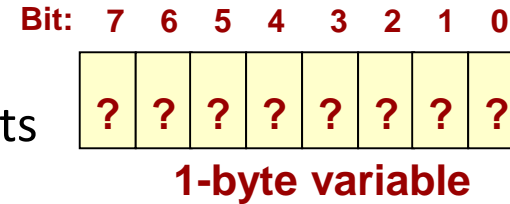
Main Point: Signed and Unsigned Multiplication require different techniques...Thus different instructions.

BITWISE & LOGIC OPERATIONS

Modifying Individual Bits

CS:APP 2.1.7

- Suppose we want to change only a single bit (or a few bits) in a variable [i.e. `char v;`] without changing the other bits
 - Set the LSB of `v` to 1 w/o affecting other bits
 - Would this work? `v = 1;`
 - Set the upper 4 bits of `v` to 1111 w/o affecting other bits
 - Would this work? `v = 0xf0;`
 - Clear the lower 2 bits of `v` to 00 w/o affecting other bits
 - Would this work? `v = 0;`
 - No!!! Assignment changes ALL bits in a variable
- **Because the smallest unit of data in computers is usually a byte, manipulating individual bits requires us to use BITWISE OPERATIONS.**



- **AND = &**
- **OR = |**
- **XOR = ^**
- **NOT = ~**

Using Bitwise Ops to Change Bits

- **ANDs** can be used to **clear a bit** (make it '0') or leave it unchanged
- **ORs** can be used to **set a bit** (make it '1') or leave it unchanged
- **XORs** can be used to **invert a bit** (flip it) or leave it unchanged

X	Y	AND
0	0	0
0	1	0
1	0	0
1	1	1

Force '0'
 Pass

0 AND y = 0
 1 AND y = y
 y AND y = y

X	Y	OR
0	0	0
0	1	1
1	0	1
1	1	1

Pass
 Force '1'

0 OR y = y
 1 OR y = 1
 y OR y = 1

X	Y	XOR
0	0	0
0	1	1
1	0	1
1	1	0

Invert
 Pass

0 XOR y = y
 1 XOR y = NOT y
 y XOR y = 0

Identity	0 OR Y = Y	1 AND Y = Y
Null Ops	1 OR Y = 1	0 AND Y = 0
Idempotency	Y OR Y = Y	Y AND Y = Y

Bitwise Operations

CS:APP 2.1.7

- The C AND , OR, XOR, NOT bitwise operations perform the operation on each pair of bits of 2 numbers

	→						
0xa5				1010	0101		
<u>AND</u> 0xf0			<u>&</u>	1111	0000		
0x30	←			1010	0000		

	→						
0xa5				1010	0101		
<u>OR</u> 0xf0			<u> </u>	1111	0000		
0xfc	←			1111	0101		

	→						
0xa5				1010	0101		
<u>XOR</u> 0xf0			<u>^</u>	1111	0000		
0x55	←			0101	0101		

	→						
<u>NOT</u> 0xa5			<u>~</u>	1010	0101		
0x5a	←			0101	1010		

```
#include <stdio.h> // C-Library
                        // for printf()

int main()
{
    char a = 0xa5;
    char b = 0xf0;

    printf("a & b = %x\n", a & b);
    printf("a | b = %x\n", a | b);
    printf("a ^ b = %x\n", a ^ b);
    printf("~a = %x\n", ~a);
    return 0;
}
```

C bitwise operators:

& = AND

| = OR

^ = XOR

~ = NOT

Logical vs. Bitwise Operations

CS:APP 2.1.8

- The C language has two types of logic operations
 - Logical and Bitwise
- Logical Operators (&&, ||, !)
- Interpret entire value as either True (non-zero) or False (zero)
- Bitwise Operators (&, |, ^, ~)
- Applies the logical operation on each pair of bits of the inputs

```
#include <stdio.h>
int main()
{
    int x = 1, y = 2;
    int z1 = x && y;
    int z2 = x & y;
    printf("z1=%d, z2=%d\n", z1, z2);

    char x = 1;
    if( !x ) { printf("L1\n"); }
    if( ~x ) { printf("L2\n"); }
    return 0;
}
```

$$\begin{array}{r} 0000 \ 0001=T \\ \&\& \ 0000 \ 0010=T \\ \hline 0000 \ 0001=T \end{array}$$

$$\begin{array}{r} 0000 \ 0001 \\ \& \ 0000 \ 0010 \\ \hline F = 0000 \ 0000 \end{array}$$

$$\begin{array}{r} ! \ 0000 \ 0001=T \\ \hline 0000 \ 0000=F \end{array}$$

$$\begin{array}{r} \sim \ 0000 \ 0001 \\ \hline T = 1111 \ 1110 \end{array}$$

$$\begin{array}{r} !! \ 0101 \ 0111=T \\ \hline 0000 \ 0001=T \end{array}$$

Important Note: Since !(non-zero) = 0; and !0 = 1
 So !!35=1. And !!-109=1

Application: Swapping via XORs

- Swapping variables can be done with a 3rd 'temp' variable
- For bitwise swapping, XORs can be used

```
#include <stdio.h>
int main()
{
    int x = 0x59, y = 0xd3;
    int temp = x;
    x = y;
    y = temp;

    return 0;
}
```

Traditional swap with 'temp'

XOR swap

```
#include <stdio.h>
int main()
{
    int x = 0x59, y = 0xd3;
    x = x ^ y;
    y = x ^ y;
    x = x ^ y;

    return 0;
}
```

$$0101\ 1001=x$$

$$1101\ 0011=y$$

$$\begin{array}{r} 0101\ 1001=x \\ \wedge\ 1101\ 0011=y \\ \hline 1000\ 1010=x \end{array}$$

$$\begin{array}{r} 1000\ 1010=x \\ \wedge\ 1101\ 0011=y \\ \hline 0101\ 1001=y \end{array}$$

$$\begin{array}{r} 1000\ 1010=x \\ \wedge\ 0101\ 1001=y \\ \hline 1101\ 0011=x \end{array}$$

Exercises

- Determine if an integer is odd (w/o % operator).
- Determine if an integer is a multiple of 4 (w/o % operator).

```
bool isOdd(int x)
{
    /* Isolate the lowest bit */
    return x&1;
}
```

```
bool isMultOf4(int x)
{
    /* Check if 2 LSBs are both 0 */
    return !(x&3);
}
```

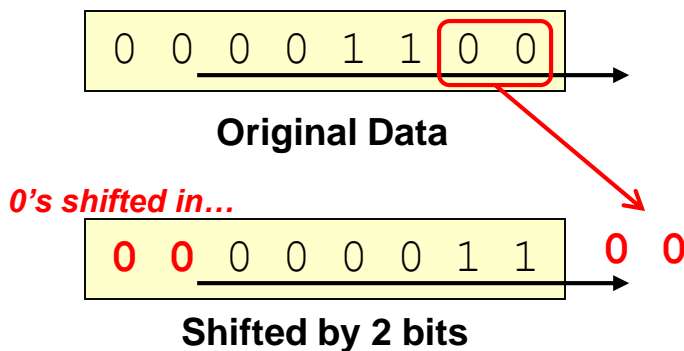
Arithmetic and Logical Shifts

SHIFT OPERATIONS

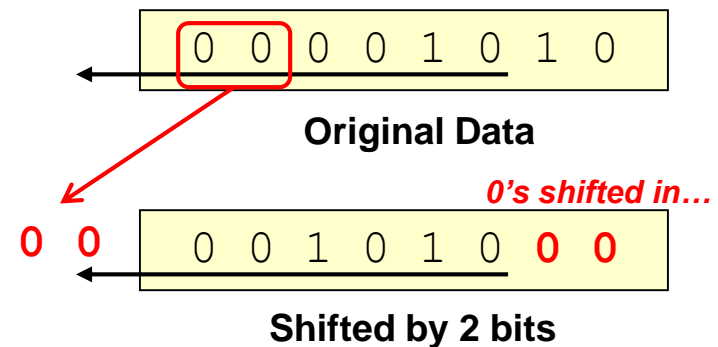
Shift Operations

- Shifts data bits either left or right
 - Bits shifted out and dropped on one side
 - Usually (but not always) 0's are shifted in on the other side
- Shifting is equivalent to multiplying or dividing by powers of 2
- 2 kinds of shifts
 - Logical shifts (used for unsigned numbers)
 - Arithmetic shifts (used for signed numbers)

Right Shift by 2 bits:



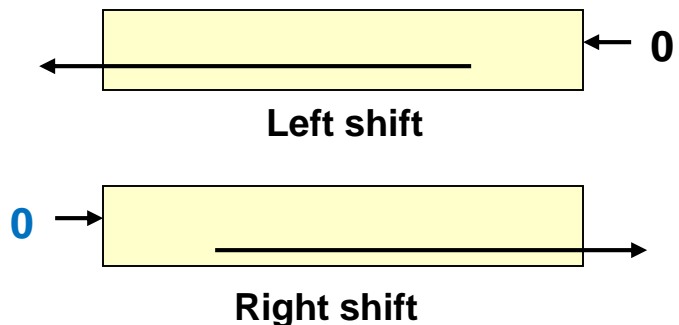
Left Shift by 2 bits:



Logical Shift vs. Arithmetic Shift

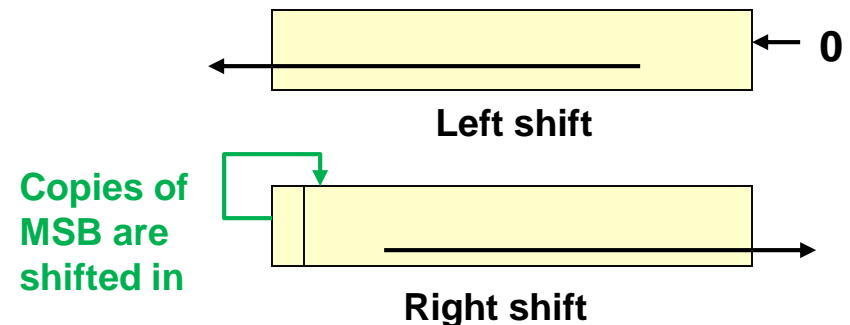
- Logical Shift

- Use for unsigned or non-numeric data
- Will always shift in 0's whether it be a left or right shift



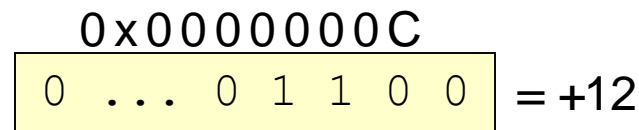
- Arithmetic Shift

- Use for signed data
- Left shift will shift in 0's
- Right shift will sign extend (replicate the sign bit) rather than shift in 0's
 - If negative number...stays negative by shifting in 1's
 - If positive...stays positive by shifting in 0's

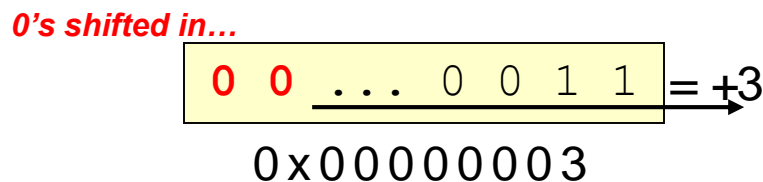


Logical Shift

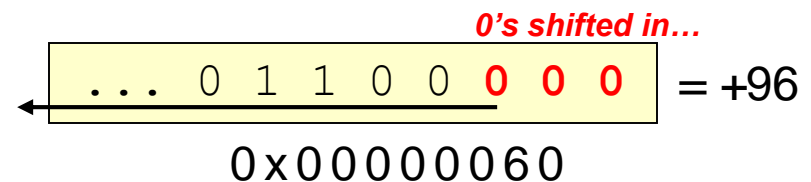
- 0's shifted in
- Only use for operations on *unsigned* data
 - Right shift by n-bits = Dividing by 2^n
 - Left shift by n-bits = Multiplying by 2^n



Logical Right Shift by 2 bits:

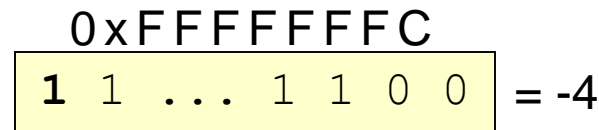


Logical Left Shift by 3 bits:



Arithmetic Shift

- Use for operations on *signed* data
- Arithmetic Right Shift – replicate MSB
 - Right shift by n-bits = Dividing by 2^n
- Arithmetic Left Shift – shifts in 0's
 - Left shift by n-bits = Multiplying by 2^n

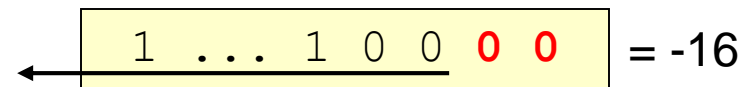
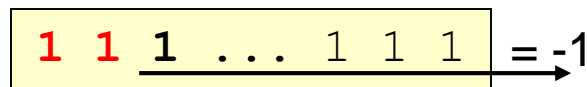


Arithmetic Right Shift by 2 bits:

Arithmetic Left Shift by 2 bits:

MSB replicated and shifted in...

0's shifted in...



0xFFFFFFFF

0xFFFFFFFF0

Notice if we shifted in 0's (like a logical right shift) our result would be a positive number and the division wouldn't work

Notice there is no difference between an arithmetic and logical left shift. We always shift in 0's.

Multiplying by Non-Powers of 2

CS:APP 2.3.6

- Left shifting by n-bits allow us to multiply by 2^n
- But what if I have to multiply a number by a *non-power* of 2 (i.e. $17*x$). Can we still use shifting?
 - Yes. Break constant into a sum using power of 2 coefficients
 - $17x = 16x + 1x$
- Exercise: How many adds/shift would be needed to compute $14*x$
 - $8x + 4x + 2x = 3$ shifts, 2 adds OR
 - $16x - 2x = 2$ shift and 1 add

$$17 = \frac{1}{16} + \frac{0}{8} + \frac{0}{4} + \frac{0}{2} + \frac{1}{1}_2$$

```
int mul17(int x)
{
    return 17*x;
}
```

Written Code

```
sall    $4, %edx
addl    %edx, %eax
```

```
int mul17(int x)
{
    int x16 = x << 4;
    return x16 + x;
}
```

Optimized Assembly
(Equivalent C)

Compiler will determine when shifts / adds become faster than constant multiplication

Integer Division By Shifting

CS:APP 2.3.7

- What is $5/2$?

- +2

- Is $5/2 = (5 \gg 1)$

- Yes

$$5 = \begin{array}{cccc} \boxed{0} & \boxed{1} & \boxed{0} & \boxed{1} \\ \hline -8 & 4 & 2 & 1 \end{array}$$

$$5 \gg 1 = \begin{array}{cccc} \boxed{0} & \boxed{0} & \boxed{1} & \boxed{0} \\ \hline -8 & 4 & 2 & 1 \end{array} \quad \begin{array}{c} \boxed{1} \\ \hline 0.5 \end{array}$$



- What is $-5/2$?

- 2

- Is $-5/2 = (-5 \gg 1)$

- No

$$-5 = \begin{array}{cccc} \boxed{1} & \boxed{0} & \boxed{1} & \boxed{1} \\ \hline -8 & 4 & 2 & 1 \end{array}$$

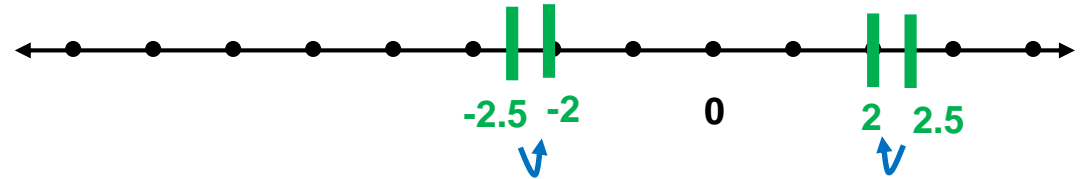
$$-5 \gg 1 = \begin{array}{cccc} \boxed{1} & \boxed{1} & \boxed{0} & \boxed{1} \\ \hline -8 & 4 & 2 & 1 \end{array} \quad \begin{array}{c} \boxed{1} \\ \hline 0.5 \end{array}$$



Main Point: Rounding fails when using shifting to divide a negative number.

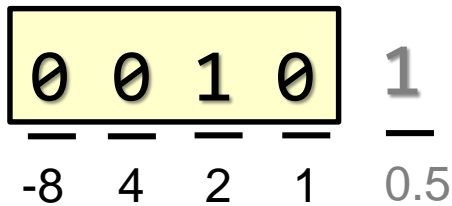
Dividing Negative Numbers

Traditional integer rounding

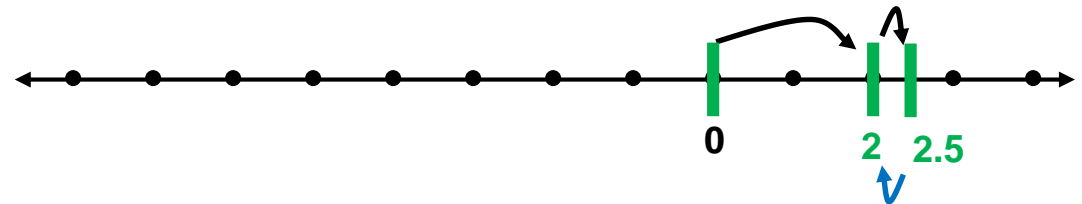


Traditional integer division rounds toward 0 (i.e. drops fractional portion)

$$+5 \gg 1$$

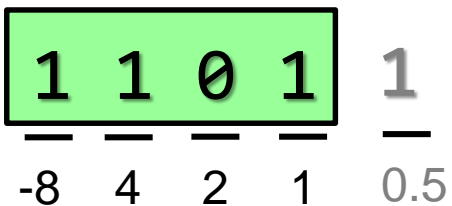


$$5 \gg 1$$

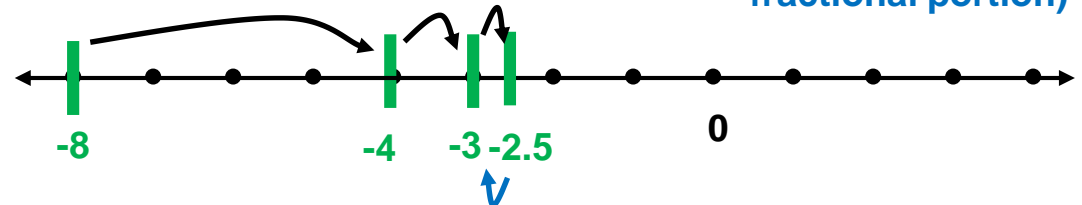


Rounding (by dropping fractional portion)

$$-5 \gg 1$$



$$-5 \gg 1$$



Rounding (by dropping fractional portion)

Main Point: Dividing numbers in the 2's complement system causes rounding to the next smallest integer, not toward 0 as desired.

Biasing

- Summary: Dividing $x / 2^k$ by performing $(x \gg k)$...
 - Works when $x \geq 0$ OR when $x < 0$ & x is a multiple of 2^k
 - Doesn't work when $x < 0$ and x is NOT a multiple of 2^k
- Idea to solve the problem:
 - Add some value (aka a **bias** value) to x before shifting that will correct for the rounding issue
 - Add $2^k - 1$ (i.e. k ones)

$$\begin{array}{r}
 -4 = 1\ 1\ 0\ 0 \\
 -4 \gg 1 = 1\ 1\ 1\ 0 \quad -2 \\
 \\
 -5 = 1\ 0\ 1\ 1 \\
 -5 \gg 1 = 1\ 1\ 0\ 1 \quad -3 \\
 \\
 \begin{array}{r}
 -5 \qquad 1\ 0\ 1\ 1 \\
 \underline{+1} \qquad + \qquad \underline{1} \\
 -4 \qquad 1\ 1\ 0\ 0 \\
 -4 \gg 1 = 1\ 1\ 1\ 0 \quad -2
 \end{array}
 \end{array}$$

More Examples

- $-8 / 4 = (-8 \ggg 2)$
 - Bias by $2^2-1 = 3$
 - $(-8 + 3) \ggg 2$

$$\begin{array}{r}
 -8 = 1\ 0\ 0\ 0 \\
 -8 \ggg 2 = 1\ 1\ 1\ 0 \quad -2
 \end{array}$$

$$\begin{array}{r}
 -8 \quad 1\ 0\ 0\ 0 \\
 +3 \quad + \quad 1\ 1 \\
 \hline
 -5 \quad 1\ 0\ 1\ 1
 \end{array}$$

$$-5 \ggg 2 = 1\ 1\ 0\ 0 \quad -2$$

- $-7 / 4 = (-7 \ggg 2)$
 - Bias by $2^2-1 = 3$
 - $(-7 + 3) \ggg 2$

$$\begin{array}{r}
 -7 = 1\ 0\ 0\ 1 \\
 -7 \ggg 2 = 1\ 1\ 0\ 0 \quad -2
 \end{array}$$

- $-20 / 16 = (-20 \ggg 4)$
 - Bias by $2^4-1 = 15$
 - $(-20 + 15) \ggg 4$

$$\begin{array}{r}
 -7 \quad 1\ 0\ 0\ 1 \\
 +3 \quad + \quad 1\ 1 \\
 \hline
 -4 \quad 1\ 1\ 0\ 0 \\
 -4 \ggg 2 = 1\ 1\ 1\ 1 \quad -1
 \end{array}$$

CS:APP Practice 2.43 (tweaked)

```
#define M /* mystery number 1 */
#define N /* mystery number 2 */

int arith(int x, int y)
{
    int result = x*M + y/N;
    return result;
}

/* Translation of assembled code for
   a given value of M and N */
int optarith(int x, int y)
{
    int t = x;
    x <<= 5;
    x -= t;
    if(y < 0) y += 3;
    y >>= 2;
    return x + y;
}
```

What were M and N when the code was compiled? (M = 31, N = 4)