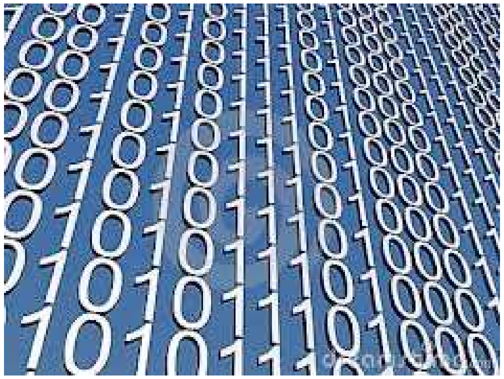
Number Systems and Number Representation



Goals of these Lectures

Help you learn (or refresh your memory) about:

- The binary, hexadecimal, and octal number systems
- Finite representation of unsigned integers
- Finite representation of signed integers
- Finite representation of rational numbers (if time)

Why?

- A power programmer must know number systems and data representation to fully understand C's **primitive data types**

Agenda

Number Systems (Lecture 1)

Finite representation of unsigned integers (Lecture 2)

Finite representation of signed integers (Lecture 3)

The Decimal Number System

Name

- "decem" (Latin) => ten

Characteristics

- Ten symbols
  - `0 1 2 3 4 5 6 7 8 9`
- Positional
  - $2945 \neq 2495$
  - $2945 = (2*10^3) + (9*10^2) + (4*10^1) + (5*10^0)$

(Most) people use the decimal number system

The Binary Number System

Name

- "binarius" (Latin) => two

Characteristics

- Two symbols
  - `0 1`
- Positional
  - $1010_B \neq 1100_B$

Most (digital) computers use the binary number system

Terminology

- **Bit**: a binary digit
- **Byte**: (typically) 8 bits

Decimal-Binary Equivalence

Decimal	Binary
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010
11	1011
12	1100
13	1101
14	1110
15	1111

Decimal	Binary
16	10000
17	10001
18	10010
19	10011
20	10100
21	10101
22	10110
23	10111
24	11000
25	11001
26	11010
27	11011
28	11100
29	11101
30	11110
31	11111
...	...

Decimal-Binary Conversion

Binary to decimal: expand using positional notation

$$
\begin{aligned}
100101_B &= (1*2^5)+(0*2^4)+(0*2^3)+(1*2^2)+(0*2^1)+(1*2^0) \\
&= 32 + 0 + 0 + 4 + 0 + 1 \\
&= 37
\end{aligned}
$$

Decimal-Binary Conversion

Decimal to binary: do the reverse

- Determine largest power of 2 ≤ number; write template

$$37 = (?*2^5)+(?*2^4)+(?*2^3)+(?*2^2)+(?*2^1)+(?*2^0)$$

- Fill in template

$$37 = (1*2^5)+(0*2^4)+(0*2^3)+(1*2^2)+(0*2^1)+(1*2^0)$$

```
 37
-32
  5
 -4
  1
 -1
  0
```

100101_B

Decimal-Binary Conversion

Decimal to binary shortcut

- Repeatedly divide by 2, consider remainder

```
37 / 2 = 18 R 1
18 / 2 =  9 R 0
 9 / 2 =  4 R 1
 4 / 2 =  2 R 0
 2 / 2 =  1 R 0
 1 / 2 =  0 R 1
```

Read from bottom to top: 100101_B

The Hexadecimal Number System

Name

- “hexa” (Greek) => six
- “decem” (Latin) => ten

Characteristics

- Sixteen symbols
  - **0 1 2 3 4 5 6 7 8 9 A B C D E F**
- Positional
  - $\mathbf{A13D_H \neq 3DA1_H}$

Computer programmers often use the hexadecimal number system

Decimal-Hexadecimal Equivalence

Decimal	Hex
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	A
11	B
12	C
13	D
14	E
15	F

Decimal	Hex
16	10
17	11
18	12
19	13
20	14
21	15
22	16
23	17
24	18
25	19
26	1A
27	1B
28	1C
29	1D
30	1E
31	1F

Decimal	Hex
32	20
33	21
34	22
35	23
36	24
37	25
38	26
39	27
40	28
41	29
42	2A
43	2B
44	2C
45	2D
46	2E
47	2F
...	...

Decimal-Hexadecimal Conversion

Hexadecimal to decimal: expand using positional notation

$$25_H = (2*16^1) + (5*16^0)$$
$$= 32 + 5$$
$$= 37$$

Decimal to hexadecimal: use the shortcut

```
37 / 16 = 2 R 5
 2 / 16 = 0 R 2
```

Read from bottom to top: 25_H

Binary-Hexadecimal Conversion

Observation: $16^1 = 2^4$

- Every 1 hexadecimal digit corresponds to 4 binary digits

Binary to hexadecimal

```
1010000100111101B
   A   1   3   DH
```

Digit count in binary number not a multiple of 4 => pad with zeros on left

Hexadecimal to binary

```
   A   1   3   DH
1010000100111101B
```

Discard leading zeros from binary number if appropriate

The Octal Number System

Name

- "octo" (Latin) => eight

Characteristics

- Eight symbols
  - **0 1 2 3 4 5 6 7**
- Positional
  - $1743_O \neq 7314_O$

Computer programmers often use the octal number system

Decimal-Octal Equivalence

Decimal	Octal
0	0
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	10
9	11
10	12
11	13
12	14
13	15
14	16
15	17

Decimal	Octal
16	20
17	21
18	22
19	23
20	24
21	25
22	26
23	27
24	30
25	31
26	32
27	33
28	34
29	35
30	36
31	37

Decimal	Octal
32	40
33	41
34	42
35	43
36	44
37	45
38	46
39	47
40	50
41	51
42	52
43	53
44	54
45	55
46	56
47	57
...	...

Decimal-Octal Conversion

Octal to decimal: expand using positional notation

$$
\begin{aligned}
37_O &= (3*8^1) + (7*8^0) \\
&= 24 + 7 \\
&= 31
\end{aligned}
$$

Decimal to octal: use the shortcut

```
31 / 8 = 3 R 7
 3 / 8 = 0 R 3
```

↑ Read from bottom to top: 37_O

Binary-Octal Conversion

Observation: $8^1 = 2^3$

- Every 1 octal digit corresponds to 3 binary digits

Binary to octal

```
001010000100111101B
 1  2  0  4  7  5O
```

Digit count in binary number not a multiple of 3 => pad with zeros on left

Octal to binary

```
 1  2  0  4  7  5O
001010000100111101B
```

Discard leading zeros from binary number if appropriate

Agenda

Number Systems (Lecture 1)

Finite representation of unsigned integers (Lecture 2)

Finite representation of signed integers (Lecture 3)

Bitwise Operations

Bitwise AND

- Similar to logical AND (`&&`), except it works on a bit-by-bit manner
- Denoted by a single ampersand: `&`

```
(1001 &
 0101)=
 0001
```

Bitwise OR

- Similar to logical OR (||), except it works on a bit-by-bit manner
- Denoted by a single pipe character: |

```
(1001 |
 0101)=
 1101
```

Bitwise XOR

- Exclusive OR, denoted by a carat: ^
- Similar to bitwise OR, except that if both inputs are 1 or 0 then the result is 0

```
(1001 ^
 0101)=
 1100
```

Bitwise NOT

- Similar to logical NOT (!), except it works on a bit-by-bit manner
- Denoted by a tilde character: ~

```
~1001 =
 0110
```

Unsigned Data Types: Java vs. C

Java has type

- **`int`**
  - Can represent signed integers

C has type:

- **`signed int`**
  - Can represent signed integers
- **`int`**
  - Same as **`signed int`**
- **`unsigned int`**
  - Can represent only unsigned integers

To understand C, must consider representation of both unsigned and signed integers

Representing Unsigned Integers

Mathematics

- Range is 0 to ∞

Computer programming

- Range limited by computer's **word** size
- Word size is n bits => range is 0 to $2^n - 1$
- Exceed range => **overflow**

Nobel computers with gcc217

- n = 32, so range is 0 to $2^{32} - 1$ (4,294,967,295)

Pretend computer

- n = 4, so range is 0 to $2^4 - 1$ (15)

Hereafter, assume word size = 4

- All points generalize to word size = 32, word size = n

Representing Unsigned Integers

On pretend computer

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

Adding Unsigned Integers

Addition

```
              1
   3         0011B
+ 10       + 1010B
  --         ----
  13         1101B
```

Start at right column

Proceed leftward

Carry 1 when necessary

```
             11
   7         0111B
+ 10       + 1010B
  --         ----
   1        10001B
```

Beware of overflow

Results are mod 2^4

Subtracting Unsigned Integers

Subtraction

```
              12
             0202
   10        1010B
-   7      - 0111B
   --        ----
    3        0011B
```

Start at right column

Proceed leftward

Borrow 2 when necessary

```
             2
    3        0011B
 - 10      - 1010B
   --        ----
    9        1001B
```

Beware of overflow

Results are mod 2^4

Shift Left

- Move all the bits `N` positions to the left, subbing in `N` `0`s on the right

Shift Left

- Move all the bits `N` positions to the left, subbing in `N` `0`s on the right

```
1001
```

Shift Left

- Move all the bits `N` positions to the left, subbing in `N` `0`s on the right

```
1001 << 2 =
100100
```

Shift Left

- Useful as a restricted form of multiplication
- Question: how?

```
1001 << 2 =
100100
```

Shift Left as Multiplication

- Equivalent decimal operation:

```
234
```

Shift Left as Multiplication

- Equivalent decimal operation:

```
234 << 1 =
2340
```

Shift Left as Multiplication

- Equivalent decimal operation:

```
234 << 1 =
2340

234 << 2 =
23400
```

Multiplication

- Shifting left `N` positions multiplies by $(\texttt{base})^{\texttt{N}}$
- Multiplying by 2 or 4 is often necessary (shift left 1 or 2 positions, respectively)
- Often a whooole lot faster than telling the processor to multiply

```
234 << 2 =
23400
```

Shift Right

- Move all the bits `N` positions to the right, subbing in **either** `N` `0`s or `N` `1`s on the left
    - Two different forms

Shift Right

- Move all the bits `N` positions to the right, subbing in **either** `N` `0`s or `N` (whatever the leftmost bit is)s on the left
  - Two different forms

```
1001 >> 2 =
either 0010 or 1110
```

Shift Right as Division

- Question: If shifting left multiplies, what does shift right do?
  - Answer: divides in a similar way, but truncates result

Shift Right as Division

- Question: If shifting left multiplies, what does shift right do?
  - Answer: divides in a similar way, but truncates result

```
234
```

Shift Right as Division

- Question: If shifting left multiplies, what does shift right do?
  - Answer: divides in a similar way, but truncates result

```
234 >> 1 =
23
```

Shifting Unsigned Integers

Bitwise right shift (>>): fill on left with zeros

```
10 >> 1 => 5
```

1010_B 0101_B

```
10 >> 2 => 2
```

1010_B 0010_B

What is the effect arithmetically? (No fair looking ahead)

Bitwise left shift (<<): fill on right with zeros

```
5 << 1 => 10
```

0101_B 1010_B

```
3 << 2 => 12
```

0011_B 1100_B

What is the effect arithmetically? (No fair looking ahead)

Results are mod 2^4

Other Operations on Unsigned Ints

Bitwise NOT (~)

- Flip each bit

```
~10 => 5
```

1010_B 0101_B

Bitwise AND (&)

- Logical AND corresponding bits

```
 10        1010B
& 7      & 0111B
 --        ----
  2        0010B
```

Useful for setting selected bits to 0

Other Operations on Unsigned Ints

Bitwise OR: (|)

- Logical OR corresponding bits

```
  10       1010B
|  1     | 0001B
  --       ----
  11       1011B
```

Useful for setting selected bits to 1

Bitwise exclusive OR (^)

- Logical exclusive OR corresponding bits

```
  10       1010B
^ 10     ^ 1010B
  --       ----
   0       0000B
```

x ^ x sets all bits to 0

The binary **XOR** operation will always produce a **1** output if either of its inputs is **1** and will produce a **0** output if both of its inputs are **0** or **1**.

Aside: Using Bitwise Ops for Arith

Can use <<, >>, and & to do some arithmetic efficiently

`x * 2^y == x << y`

- `3*4 = 3*2^2 = 3 << 2 => 12`

  0011_B → 1100_B

Fast way to **multiply** by a power of 2

`x / 2^y == x >> y`

- `13/4 = 13/2^2 = 13 >> 2 => 3`

  1101_B → 0011_B

Fast way to **divide** by a power of 2

`x % 2^y == x & (2^y-1)`

- `13%4 = 13%2^2 = 13&(2^2-1)`
  `= 13&3 => 1`

Fast way to **mod** by a power of 2

```
 13        1101B
& 3      & 0011B
--         ----
  1        0001B
```

Two Forms of Shift Right

- Subbing in `0`s makes sense
- What about subbing in the leftmost bit?
  - And why is this called "arithmetic" shift right?

```
1100 (arithmetic)>> 1 =
1110
```

Answer... Sort of

- Arithmetic form is intended for numbers in *two*'s *complement* (next lecture), whereas the non-arithmetic form is intended for *unsigned* numbers

Agenda

Number Systems (Lecture 1)

Finite representation of unsigned integers (Lecture 2)

Finite representation of signed integers (Lecture 3)

Signed Magnitude

Integer	Rep
-7	1111
-6	1110
-5	1101
-4	1100
-3	1011
-2	1010
-1	1001
-0	1000
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111

Definition

High-order bit indicates sign

- 0 => positive
- 1 => negative

Remaining bits indicate magnitude

$1101_B = -101_B = -5$

$0101_B = 101_B = 5$

Signed Magnitude (cont.)

Integer	Rep
-7	1111
-6	1110
-5	1101
-4	1100
-3	1011
-2	1010
-1	1001
-0	1000
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111

Computing negative

neg(x) = flip high order bit of x

$neg(0101_B) = 1101_B$

$neg(1101_B) = 0101_B$

Pros and cons

+ easy for people to understand
+ symmetric
- two reps of zero
- one of the bit patterns is wasted.
- addition doesn't work the way we want it to.

Signed Magnitude (cont.)

Problem #1: "The Case of the Missing Bit Pattern":

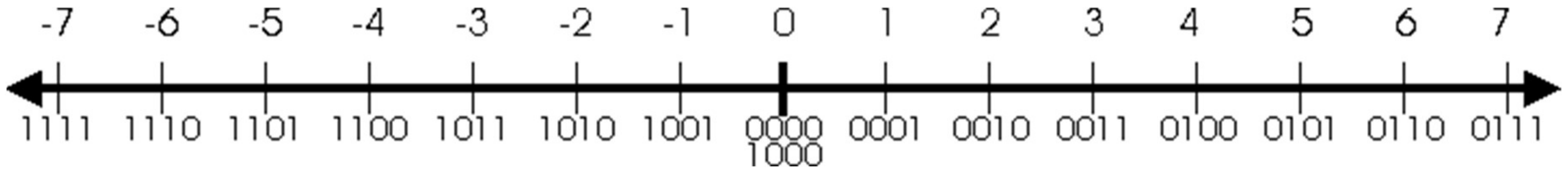
How many possible bit patterns can be created with 4 bits?

Easy, we know that's 16. In unsigned representation, we were able to represent 16 numbers: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15.

But with signed magnitude, we are only able to represent 15 numbers: -7, -6, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, 6, and 7.

There's still 16 bit patterns, but one of them is either not being used or is duplicating a number. That bit pattern is '1000B'.

When we interpret this pattern, we get '-0' which is both nonsensical (negative zero? come on!) and redundant (we already have '0000B' to represent 0).



Signed Magnitude (cont.)

Problem #2: "Requires Special Care and Feeding": Remember we wanted to have negative binary numbers so we could use our binary addition algorithm to simulate binary subtraction. How does signed magnitude fare with addition? To test it, let's try subtracting 2 from 5 by adding 5 and -2. A positive 5 would be represented with the bit pattern '0101B' and -2 with '1010B'. Let's add these two numbers and see what the result is:

```
 0101
+1010
----------
 1111
```

Now we interpret the result as a signed magnitude number. The sign is ‘1’ (negative) and the magnitude is '7'. So the answer is a negative 7. But, wait a minute, 5-2=3! This obviously didn't work.

Conclusion: signed magnitude doesn't work with regular binary addition algorithms.

One's Complement

Integer	Rep
-7	1000
-6	1001
-5	1010
-4	1011
-3	1100
-2	1101
-1	1110
-0	1111
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111

Definition

High-order bit has weight -7 ($-2^n + 1$)

$1010_B = (1*-7)+(0*4)+(1*2)+(0*1)$
$= -5$

$0010_B = (0*-7)+(0*4)+(1*2)+(0*1)$
$= 2$

One's Complement (cont.)

Integer	Rep
-7	1000
-6	1001
-5	1010
-4	1011
-3	1100
-2	1101
-1	1110
-0	1111
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110

Computing negative

neg(x) = ~x

$neg(0101_B) = 1010_B$

$neg(1010_B) = 0101_B$

Computing negative (alternative)

$neg(x) = 1111_B - x$

$neg(0101_B) = 1111_B - 0101_B$

$= 1010_B$

$neg(1010_B) = 1111_B - 1010_B$

$= 0101_B$

Pros and cons

+ symmetric

- two reps of zero

Two's Complement

Integer	Rep
-8	1000
-7	1001
-6	1010
-5	1011
-4	1100
-3	1101
-2	1110
-1	1111
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111

Definition

High-order bit has weight -8 (-2^n)

$$1010_B = (1*-8)+(0*4)+(1*2)+(0*1) = -6$$

$$0010_B = (0*-8)+(0*4)+(1*2)+(0*1) = 2$$

Two's Complement (cont.)

Integer	Rep
-8	1000
-7	1001
-6	1010
-5	1011
-4	1100
-3	1101
-2	1110
-1	1111
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111

Computing negative

neg(x) = ~x + 1

neg(x) = onescomp(x) + 1

$\mathtt{neg}(0101_B) = 1010_B + 1 = 1011_B$

$\mathtt{neg}(1011_B) = 0100_B + 1 = 0101_B$

Pros and cons

- not symmetric

+ one rep of zero

Two’s Complement (cont.)

Almost all computers use two’s complement to represent signed integers

Why?

- Arithmetic is easy
- Will become clear soon

Hereafter, assume two’s complement representation of signed integers

Two's Complement

- Way to represent positive integers, negative integers, and zero
- If `1` is in the *most significant bit* (generally leftmost bit in this class), then it is negative

Decimal to Two's Complement

- Example: -5 decimal to binary (twos complement)

http://sandbox.mc.edu/~bennet/cs110/tc/dtotc.html

Decimal to Two's Complement

- Example: -5 decimal to binary (twos complement)
- First, convert the magnitude to an unsigned representation

Decimal to Two's Complement

- Example: -5 decimal to binary (two's complement)
- First, convert the magnitude to an unsigned representation

5 (decimal) = 0101 (binary)

Decimal to Two's Complement

- Then, take the bits, and negate them

Decimal to Two's Complement

- Then, take the bits, and negate them

```
0101
```

Decimal to Two's Complement

- Then, take the bits, and negate them

```
~0101 =
 1010
```

Decimal to Two's Complement

- Finally, add one:

Decimal to Two's Complement

- Finally, add one:

```
1010
```

Decimal to Two's Complement

- Finally, add one:

```
1010 + 1 =
1011
```

Two's Complement to Decimal

- Same operation: negate the bits, and add one

Two's Complement to Decimal

- Same operation: negate the bits, and add one

```
1011
```

Two's Complement to Decimal

- Same operation: negate the bits, and add one

```
~1011 =
 0100
```

Two's Complement to Decimal

- Same operation: negate the bits, and add one

```
0100
```

Two's Complement to Decimal

- Same operation: negate the bits, and add one

```
0100 + 1 =
0101
```

Two's Complement to Decimal

- Same operation: negate the bits, and add one

```
0100 + 1 =
0101 =
-5
```

We started with `1011` - negative

Addition

http://sandbox.mc.edu/~bennet/cs110/textbook/module3_2.html

Building Up Addition

- Question: how might we add the following, in decimal?

```
  986
+123
----
    ?
```

Building Up Addition

- Question: how might we add the following, in decimal?

```
  986
+123
----
    ?
```

			6 +3 -- ?

Building Up Addition

- Question: how might we add the following, in decimal?

```
  986
+123
----
    ?
```

		`  8`
		`+2`
		`--`
		`  ?`
			`  6`
			`+3`
			`--`
			`  9`

Building Up Addition

- Question: how might we add the following, in decimal?

```
  986
+123
----
    ?
```

	Carry: 1		
		8 +2 -- 0	6 +3 -- 9

Building Up Addition

- Question: how might we add the following, in decimal?

```
  986
+123
----
    ?
```

	1 9 +1 -- ?	8 +2 -- 0	6 +3 -- 9

Building Up Addition

- Question: how might we add the following, in decimal?

```
  986
+123
----
    ?
```

Carry: 1	1 9 +1 -- 1	8 +2 -- 0	6 +3 -- 9

Building Up Addition

- Question: how might we add the following, in decimal?

```
  986
+123
----
    ?
```

1 +0 -- 1	1 9 +1 -- 1	8 +2 -- 0	6 +3 -- 9

Core Concepts

- We have a "primitive" notion of adding single digits, along with an idea of *carrying* digits
- We can build on this notion to add numbers together that are more than one digit long

Now in Binary

- Arguably simpler - fewer one-bit possibilities

0 +0 -- ?	0 +1 -- ?	1 +0 -- ?	1 +1 -- ?

Now in Binary

- Arguably simpler - fewer one-bit possibilities

0	0	1	1
+0	+1	+0	+1
--	--	--	--
0	1	1	0
			Carry: 1

Chaining the Carry

- Also need to account for any input carry

0 0 +0 -- 0	0 0 +1 -- 1	0 1 +0 -- 1	0 1 +1 -- 0 Carry: 1
1 0 +0 -- 1	1 0 +1 -- 0 Carry: 1	1 1 +0 -- 0 Carry: 1	1 1 +1 -- 1 Carry: 1

Adding Multiple Bits

- How might we add the numbers below?

```
  011
 +001
------
```

Adding Multiple Bits

- How might we add the numbers below?

```
   0
 011
+001
------
```

Adding Multiple Bits

- How might we add the numbers below?

```
  10
 011
+001
------
   0
```

Adding Multiple Bits

- How might we add the numbers below?

```
  110
  011
 +001
------
   00
```

Adding Multiple Bits

- How might we add the numbers below?

```
 0110
  011
 +001
------
  100
```

Adding Multiple Bits

- How might we add the numbers below?

```
 0110
  011
 +001
------
  100
```

Output Carry Bit

Result Bits

Another Example

```
  111
 +001
------
```

Another Example

```
      0
   111
  +001
------
```

Another Example

```
   10
  111
 +001
------
    0
```

Another Example

```
  110
  111
 +001
------
   00
```

Another Example

```
 1110
  111
 +001
------
  000
```

Output Carry Bit

Result Bits

Output Carry Bit Significance

- For unsigned numbers, it indicates if the result did not fit all the way into the number of bits allotted
- May be an error condition for software

Signed Addition

- Question: what is the result of the following operation?

```
 011
+011
----
   ?
```

Signed Addition

- Question: what is the result of the following operation?

```
 011
+011
----
0110
```

Overflow

- In this situation, *overflow* occurred: this means that both the operands had the same sign, and the result's sign differed

```
 011
+011
----
 110
```

- Possibly a software error

Overflow vs. Carry

- These are **different ideas**
  - Carry is relevant to **unsigned** values
  - Overflow is relevant to **signed** values

111 +001 ---- 000	011 +011 ---- 110	111 +100 ---- 011	001 +001 ---- 010
No Overflow; Carry	Overflow; No Carry	Overflow; Carry	No Overflow; No Carry

Adding Signed Integers

pos + pos

```
            11
  3         0011B
+ 3       + 0011B
 --         ----
  6         0110B
```

pos + pos (overflow)

```
            111
  7         0111B
+ 1       + 0001B
 --         ----
 -8         1000B
```

pos + neg

```
           1111
   3         0011B
+ -1       + 1111B
  --         ----
   2        10010B
```

neg + neg

```
            11
  -3         1101B
+ -2       + 1110B
  --         ----
  -5        11011B
```

neg + neg (overflow)

```
            1 1
  -6         1010B
+ -5       + 1011B
  --         ----
   5        10101B
```

Subtracting Signed Integers

Perform subtraction with borrows

or

Compute two's comp and add

	1
	22
3	0011_B
- 4	- 0100_B
-1	1111_B

3	0011_B
+ -4	+ 1100_B
-1	1111_B

-5	1011_B
- 2	- 0010_B
-7	1001_B

	111
-5	1011
+ -2	+ 1110
-7	11001

Shifting Signed Integers

Bitwise (**logical/arithmetic**) left shift (<<): fill on right with zeros

```
3 << 1 => 6
```

0011_B 0110_B

```
-3 << 1 => -6
```

1101_B 1010_B

Shift by n = multiplying by 2^n

Bitwise **arithmetic** right shift: fill on left **with sign bit**

```
6 >> 1 => 3
```

0110_B 0011_B

```
-6 >> 1 => -3
```

1010_B 1101_B

Shift by n = dividing by 2^n and Round-floor

Results are mod 2^4

Shifting Signed Integers (cont.)

Bitwise **logical** right shift: fill on left **with zeros**

```
6 >> 1 => 3
```

0110_B 0011_B

```
-6 >> 1 => 5
```

?

1010_B 0101_B

Right shift (>>) could be logical or arithmetic

- Compiler designer decides
- **Logical** shift is ideal for unsigned binary numbers
- **Arithmetic** shift is ideal for signed two's complement binary numbers

Other Operations on Signed Ints

Bitwise NOT (~)

- Same as with unsigned ints

Bitwise AND (&)

- Same as with unsigned ints

Bitwise OR: (|)

- Same as with unsigned ints

Bitwise exclusive OR (^)

- Same as with unsigned ints

Bitwise Operations as Masks

X: it is an unknown binary number and can be either 0 or 1

AND (&) Operation:

X & 0 = 0 & X = **0**
X & 1 = 1 & X = X
X & X = X

OR (|) Operation:

X | 1 = 1 | X = **1**
X | 0 = 0 | X = X
X | X = X

XOR (^) Operation:

X ^ 1 = 1 ^ X = ~X
X ^ 0 = 0 ^ X = X
X ^ X = 0

Mask Example

Specify the mask you would need to isolate bit 0 of the unknown number. The result of the operation should be **0 (0x0000) if bit 0 is 0, and non-zero if bit 0 is 1**. Express it as a 4-digit hexadecimal number.

Answer:

We know that 1 hexadecimal digit = 4 bits in binary

```
                  15...              ...... 3 2 1 0   ← Bit position
                   XXXX  XXXX  XXXX XXXX              ← Unknown number
Operation --> ?    ????  ????  ????  ????             ← Mask
               -------------------------------------
if bit 0 is 0  →  0000  0000  0000  0000              ← zero (0x0000)
if bit 0 is 1  →  0000  0000  0000   0001             ← nonzero (0x0001)
```

In this case, we can use AND operation (**&**) and then the mask(16 bits) will be as

0000 0000 0000 0001 => 0001 in hexadecimal

Therefore, the answer is answer **& as the operation and 0x0001 as the mask.**

Mask Example

Specify the mask you would need to **set bit 1 of the unknown number to zero**. That is, the result of this operation results in a new number, which the unknown number will be subsequently set to. Express it as a 4-digit hexadecimal number.

Answer:

We know that 1 hexadecimal digit = 4 bits in binary

```
                 15...          ...... 3 2 1 0   ← Bit position
                  XXXX  XXXX  XXXX XXXX          ← Unknown number
Operation --> ?   ????  ????  ????  ????         ← Mask
               -----------------------------------
                  XXXX  XXXX  XXXX  XX0X
```

In this case, we can use AND operation (**&**) and then the mask(16 bits) will be as

1111 1111 1111 1101 => FFFD in hexadecimal

Therefore, the answer is **& as the operation and 0xFFFD as the mask.**

Summary

The binary, hexadecimal, and octal number systems

Finite representation of unsigned integers

Finite representation of signed integers

Essential for proper understanding of

- C or Java primitive data types
- Assembly language
- Machine language