

香港中文大學

The Chinese University of Hong Kong

CSCI2510 Computer Organization

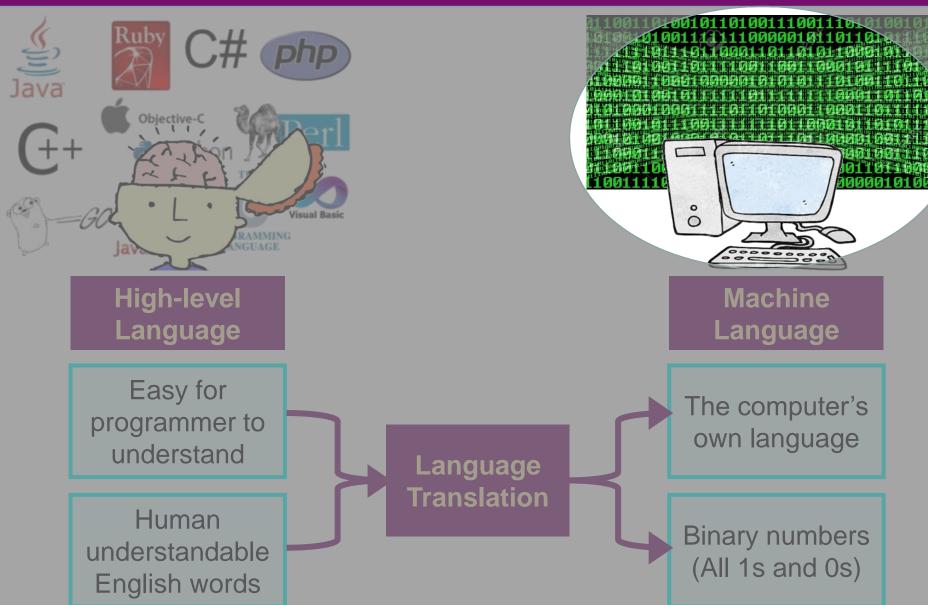
Lecture 02: Number and Character Representation

Ming-Chang YANG



Recall: How to talk to the computer?





Outline



- Number Representation
 - Number Systems
 - Integers
 - Unsigned Integer
 - Signed Integer
 - Arithmetic Operations
 - Floating-Point Numbers
 - Unsigned Binary Fraction
 - Floating-Point Number Representation
 - Arithmetic Operations
- Character Representation
 - ASCII

Number Systems

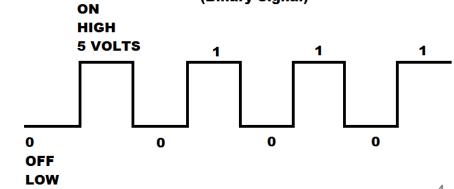


- Common number systems:
 - The radix or base of the number system denotes the number of digits used in the system.

Binary (base 2)	0	1														
Octal (base 8)	0	1	2	3	4	5	6	7								
Decimal (base 10)	0	1	2	3	4	5	6	7	8	9						
Hexadecimal (base 16)	0	1	2	3	4	5	6	7	8	9	A	В	С	D	Ε	F

0 VOLTS

- The most natural way in a computer system is by binary numbers (0, 1).
 - (0, 1) can be represented as (off, on) electrical signals.



Outline



- Number Representation
 - Number Systems
 - Integers
 - Unsigned Integer
 - Signed Integer
 - Arithmetic Operations
 - Floating-Point Numbers
 - Unsigned Binary Fraction
 - Floating-Point Number Representation
 - Arithmetic Operations
- Character Representation
 - ASCII

Count to 100 in Decimal!



1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41

the Count to 100 by Ones Song

66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

 $100 = 1 \times 10^2 + 0 \times 10^1 + 0 \times 10^0$

"Unsigned" Integer Representation



Consider an *n*-bit (or *n*-digit) vector

$$B = (b_{n-1} \dots b_1 b_0)_2$$
; as a subscript

Denoting the base

where $b_i = 0$ or 1 (binary number) for $0 \le i \le n-1$

- Most Significant Bit (MSB): b_{n-1} (i.e., the leftmost bit)
- Least Significant Bit (LSB): b_0 (i.e., the rightmost bit)
- This vector can represent the decimal value for an unsigned integer V(B) in the range 0 to $2^n - 1$, where $V(B) = b_{n-1} \times 2^{n-1} + \dots + b_1 \times 2^1 + b_0 \times 2^0$
- For example, if $B = (1001)_2$, where n = 4 $V(B) = \frac{1}{1} \times \frac{2^3}{1} + \frac{0}{1} \times \frac{2^2}{1} + \frac{0}{1} \times \frac{2^1}{1} + \frac{1}{1} \times \frac{2^0}{1} = (9)_{10}$

Conversion of Number Systems



(Decimal) ₁₀	(Binary) ₂	(Octal) ₈	(Hexadecimal) ₁₆
(00) ₁₀	(0000) ₂	(00) ₈	(0) ₁₆
(01) ₁₀	(0001) ₂	(01) ₈	(1) ₁₆
(02) ₁₀	(0010) ₂	(02) ₈	(2) ₁₆
(03) ₁₀	(0011) ₂	(03) ₈	(3) ₁₆
(04) ₁₀	(0100) ₂	(04) ₈	(4) ₁₆
(05) ₁₀	(0101) ₂	(05) ₈	(5) ₁₆
(06) ₁₀	(0110) ₂	(06) ₈	(6) ₁₆
(07) ₁₀	(0111) ₂	(07) ₈	(7) ₁₆
(08) ₁₀	(1000) ₂	(10) ₈	(8) ₁₆
(09) ₁₀	(1001) ₂	(11) ₈	(9) ₁₆
(10) ₁₀	(1010) ₂	(12) ₈	(A) ₁₆
(11) ₁₀	(1011) ₂	(13) ₈	(B) ₁₆
(12) ₁₀	(1100) ₂	(14) ₈	(C) ₁₆
(13) ₁₀	(1101) ₂	(15) ₈	(D) ₁₆
(14) ₁₀	(1110) ₂	(16) ₈	(E) ₁₆
(15) ₁₀	(1111) ₂	(17) ₈	(F) ₁₆

Class Exercise 2.1

Student ID: _____ Date: Name: ____

• Represent (255) 10 in binary, octal, and hexadecimal:

Binary (base 2)	0	1														
Octal (base 8)	0	1	2	3	4	5	6	7								
Decimal (base 10)	0	1	2	3	4	5	6	7	8	9						
Hexadecimal (base 16)	0	1	2	3	4	5	6	7	8	9	Α	В	С	D	Ε	F

"Signed" Integer Representation (1/3)



- To represent both positive and negative numbers, we need different systems to representing signed integer.
- In <u>written</u> decimal system, a signed integer is usually represented by a "+" or "-" <u>sign</u> and followed by the magnitude.
 - E.g. -73, -215, +349
- In binary system, we have three common systems:
 - ① Sign-and-magnitude
 - 2 1's-complement
 - 3 2's-complement

"Signed" Integer Representation (2/3)



- The leftmost bit (MSB) decides the sign (0: "+", 1: "-").
 - Positive values are identical in all the three systems:
 - Rule: Treating the rest bits as an unsigned integer \triangleright E.g., +3 is represented by 0011.
 - Negative values have different representations:
 - ① **Sign-and-magnitude** (MSB: sign, other bits: magnitude)
 - ex: 0011 Rule: Changing the MSB from 0 to 1 ➤ E.g. -3 is represented by 1011. 1011
 - ② 1's-complement
 - *Rule*: Inverting each bit of the positive number
 - ➤ E.g. -3 is obtained by flipping each bit in 0011 to yield 1100.
 - 3 2's-complement
 - Rule 1: Subtracting the positive number from the unsigned 2ⁿ
 - Rule 2: Adding 1 to 1's-complement of that negative number
- ➤ E.g. –3 is represented by 1101 when applying either rule. CSCI2510 Lec02: Number and Character Representation

1101

+) 0001

10000 -) 0011

1101

1100

ex:

ex:

ex: 0011

1100

"Signed" Integer Representation (3/3)



В	Values Represented in Decimal										
$b_3b_2b_1b_0$	Sign-and-magnitude	1's-complement	2's-complement								
0111	+ 7	+ 7	+ 7								
0110	+ 6	+ 6	+ 6								
0101	+ 5	+ 5	+ 5								
0100	+ 4	+ 4	+ 4								
0011	+ 3	+ 3	+ 3								
0010	+ 2	+ 2	+ 2								
0001	+ 1	+ 1	+ 1								
0 0 0 0	+ 0	+ 0	+ 0								
1000	- 0	- 7	- 8								
1001	- 1	- 6	- 7								
1010	- 2	- 5	- 6								
1011	- 3	- 4	- 5								
1100	- 4	- 3	- 4								
1101	- 5	- 2	- 3								
1110	- 6	- 1	- 2								
1111	- 7	– 0	- 1								

Class Exercise 2.2



- Question: Which representation system(s) uses distinct representations for +0 and -0?
- Answer: _____
- Question: Which representation system(s) has only one representation for 0?
- Answer: ______
- Question: Which representation system(s) is able to represent – 8 for 4-bit numbers?
- Answer: ______

Class Exercise 2.3



- Question: Consider the decimal number -56. Please use 8 bits to represent it in:
 - Sign-and-magnitude:
 - 1's-complement:
 - 2's-complement:
- Question: Consider the 8-bit string 10110101, what is its decimal value when interpreted as:
 - Sign-and-magnitude: _____
 - 1's-complement:
 - 2's-complement:
- Question: Given n bits, what is the range of integers can be represented by the three representations?
- Answer:

Addition of "Unsigned" Integers



Addition of 1-bit unsigned numbers:

carry-out sum

- To add multiple-bit numbers:
 - We add bit pairs starting from the low-order (right) end,
 propagating carries toward the high-order (left) end.
 - The carry-out from a bit pair becomes the carry-in to the next bit pair.
 - The carry-in must be added to a bit pair in generating the sum and carry-out at that position.

 carry-in 1
 - For example, 01111111 + 00000001

1000000

low-order

Arithmetic of "Signed" Integers



- The three signed integer representation systems differ only in the way of representing negative values.
- Their relative merits on performing arithmetic operations can be summarized as follows:
 - Sign-and-magnitude: the simplest representation, but it is also the most awkward for addition/subtraction operations.
 - 1's-complement: somewhat better than the sign-andmagnitude system.
 - 2's-complement: the most efficient method for performing addition and subtraction operations.
 - This is also why the 2's-complement system is the one most often used in modern computers.

Why 2's-complement Arithmetic?



- First consider adding +7 to -3:
 - What if we perform this addition by adding bit pairs from right to left (as what we did for n-bit unsigned numbers)?

- If the leftmost carry-out bit is ignored, we get (+4)₁₀.
- Rules for n-bit signed number addition/subtraction:
 - -X+Y
 - Add their n-bit 2's-complement representations from right to left
 - Ignore the carry-out bit at the MSB position
 - -X-Y
 - Interpret as, and perform X + (-Y)
 - Note: The sum should be in the range of $-2^{n-1} \sim (2^{n-1}-1)$

Class Exercise 2.4



• Using 4-bit 2's-complement number to calculate:

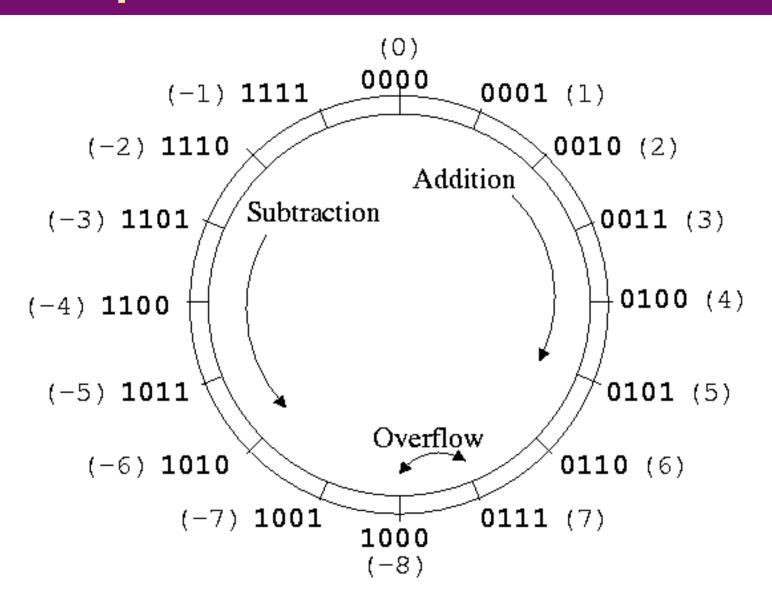
•
$$4 + (-6)$$

•
$$(-5)+(-2)$$

•
$$(-7) - (-5)$$

2's-Complement Number Wheel





Overflow in Integer Arithmetic



- Overflow: The result of an arithmetic operation does not fall within the representable range.
 - In Unsigned Number Arithmetic:
 - A carry-out of 1 from the MSB-bit always indicates an overflow.
 - E.g. $(1111)_2 + (0001)_2 = (1 0000)_2 \leftarrow overflowed$
 - E.g. $(0111)_2 + (0001)_2 = (0 \ 1000)_2 \leftarrow no \ overflow$
 - In 2's-complement Signed Number Arithmetic:
 - The carry-out bit from the sign-bit is not an indicator of overflow.
 - E.g. $(+7)_{10}$ + $(+4)_{10}$ = $(0111)_2$ + $(0100)_2$ = $(\underline{0} \ 1011)_2$ = $(-5)_{10}$
 - E.g. $(-4)_{10}$ + $(-6)_{10}$ = $(1100)_2$ + $(1010)_2$ = $(\underline{1} 0110)_2$ = $(+6)_{10}$
 - If the numbers are the same sign and the result is the opposite sign, we say that an overflow has occurred.
 - E.g. $(+7)_{10}$ + $(+4)_{10}$ = $(0111)_2$ + $(0100)_2$ = $(1011)_2$ = $(-5)_{10}$ \leftarrow overflowed
 - E.g. $(-4)_{10}$ + $(-6)_{10}$ = $(1100)_2$ + $(1010)_2$ = $(0110)_2$ = $(+6)_{10}$ \leftarrow overflowed
 - Addition of opposite sign numbers never causes overflow
 - E.g. $(+7)_{10}$ + $(-6)_{10}$ = $(0111)_2$ + $(1010)_2$ = $(0001)_2$ = $(+1)_{10}$ \leftarrow no overflow

Sign Extension



- We often need to represent a value given in a certain number of bits by using a larger number of bits.
 - That is, how to represent a signed integer by using a larger number of bits?
- Sign Extension: Simply repeat the "sign bit" as many times as needed to the left. (Note: It can be applied to both 1's and 2's-complement, but not sign-and-magnitude)
 - Positive Number: Add 0's to the left-hand-side
 - E.g. $0111 \rightarrow 0000 \ 0111$
 - Negative Number: Add 1's to the left-hand-side
 - E.g. $1010 \rightarrow 1111 \ 1010$

Example: Representing -2~+1 with 8 bits by 2's-complement

$B = b_7 b_6$	b ₀	2's complement
000000	0 1	+ 1
000000	0 0	+ 0
111111	1 0	- 2
111111	1 1	- 1

Outline



- Number Representation
 - Number Systems
 - Integers
 - Unsigned Integer
 - Signed Integer
 - Arithmetic Operations
 - Floating-Point Numbers
 - Unsigned Binary Fraction
 - Floating-Point Number Representation
 - Arithmetic Operations
- Character Representation
 - ASCII

Unsigned Binary Fraction



Consider a n-bit unsigned binary fraction:

$$B = (0.\,b_{-1}b_{-2}\,...\,b_{-n})_2$$
 where $b_{-i}=0$ or 1 (binary number) for $1\leq i\leq n$

• This vector can represent the value for an <u>unsigned</u> binary fraction F(B), where

$$F(B) = b_{-1} \times 2^{-1} + b_{-2} \times 2^{-2} + \dots + b_{-n} \times 2^{-n}$$

• The range of F(B) is

$$0 \le F(B) \approx +1.0$$
, for a large n

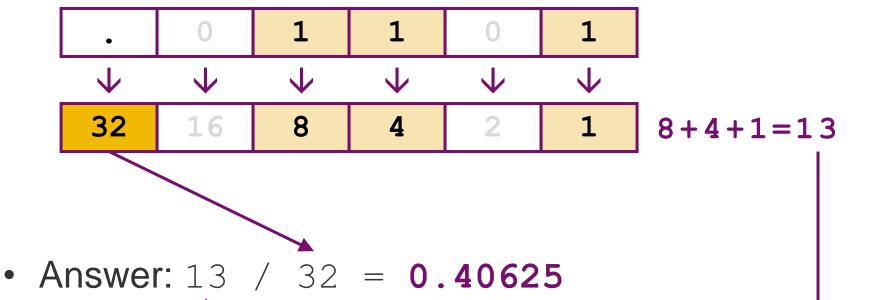
B) is
$$0 \le F(B) \le 1 - 2^{-n}$$
 $S_n = \sum_{i=1}^n a_i r^{i-1} = a_1 \left(\frac{1 - r^n}{1 - r}\right)$

Why? Geometric Series

Binary Fraction to Decimal Fraction



What is the binary fraction (0.011010)₂ in decimal?



CSCI2510 Lec02: Number and Character Representation

Decimal Fraction to Binary Fraction



What is the decimal fraction (0.6875)₁₀ in binary?

$$0.6875 * 2 = 1.3750 \rightarrow 0.1???_2$$
 $0.3750 * 2 = 0.7500 \rightarrow 0.10??_2$
 $0.7500 * 2 = 1.5000 \rightarrow 0.101?_2$
 $0.5000 * 2 = 1.0000 \rightarrow 0.1011_2$
 $0.0000 * 2 = 0 \rightarrow End$

Answer: (0.1011)₂

Why? Let's have an analogy in decimal:

$$0.6875 * 10 = 6.875 \rightarrow (0.6???)_{10}$$

 $0.8750 * 10 = 8.7500 \rightarrow (0.68??)_{10}$

• • •

Class Exercise 2.5



- What is the decimal fraction (0.1)₁₀ in binary?
- Answer:

What did we learn so far?



- Some decimal fractions (e.g. (0.1)₁₀) will produce infinite binary fraction expansions.
- The position of the binary point in a floating-point number varies (that's way called floating point!).

```
0.232 * 10^4 = 2.320000 * 10^3
= 23.20000 * 10^2
```

- A 32-bit signed integer in 2's-complement form can only represent values in the range of $-2^{31} \sim 2^{31} 1$.
- · We need a unique representation that can
 - Represent the sign, and the position of the floating point.
 - Represent both very large integers and very small fractions.

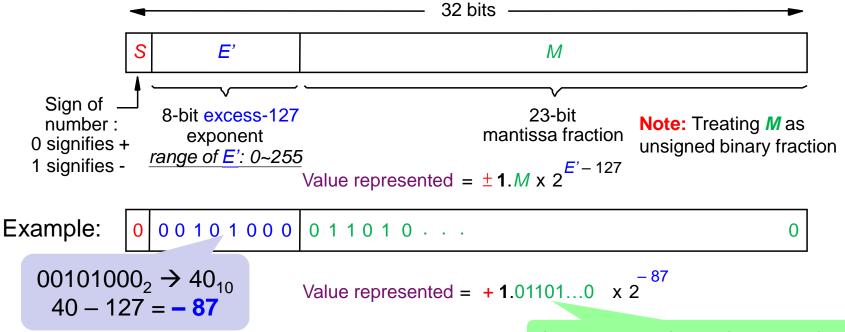
Floating Point Number Representation

- In decimal scientific notation, numbers are written as : $+6.0247 \times 10^{23}$, $+3.7291 \times 10^{-27}$, -7.3000×10^{-14} , ...
- The same approach can be used to represent binary floating-point numbers (using 2 as the base) by:
 - Sign: A sign for the number
 - Mantissa: Some significant bits
 - Exponent: A signed scale factor (implied base of 2)
- To have a normalized representation for floating-point numbers, we should normalize Mantissa in the range [1 ... B), where B is the base.
 - Binary System: [1 ... 2)
 - $(1.b_{-1}b_{-2}...b_{-n})_2$ must in the range of [1...2).

IEEE Standard 754 Single Precision



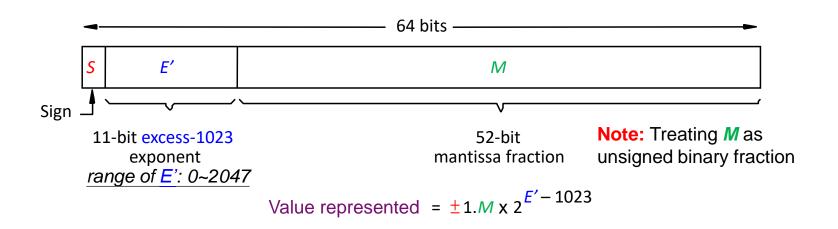
- The single precision format is a 32-bit representation.
 - The leftmost bit represents the sign, S, for the number.
 - The next 8 bits, E', represent the unsigned integer for the excess – 127 exponent (with base of 2).
 - Note: The actual signed exponent E is E'-127
 - The remaining 23 bits, M, are the significant bits.



IEEE Standard 754 Double Precision



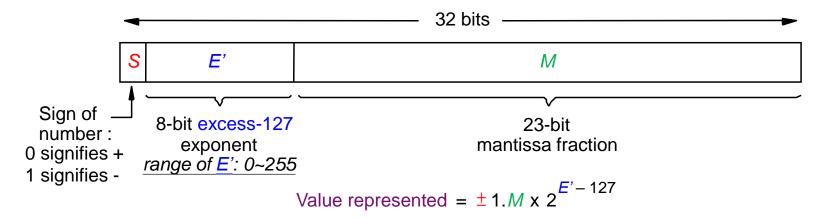
- The double precision format is a 64-bit representation.
 - The leftmost bit represents the sign, S, for the number.
 - The next 11 bits, E', represent the unsigned integer for the excess-1023 exponent (with base of 2).
 - Note: The actual signed exponent E is E'-1023.
 - The remaining 52 bits, M, are the significant bits.



Example of IEEE Single Precision



- What is the IEEE single precision number
 4000 000016 in decimal?
- Answer:

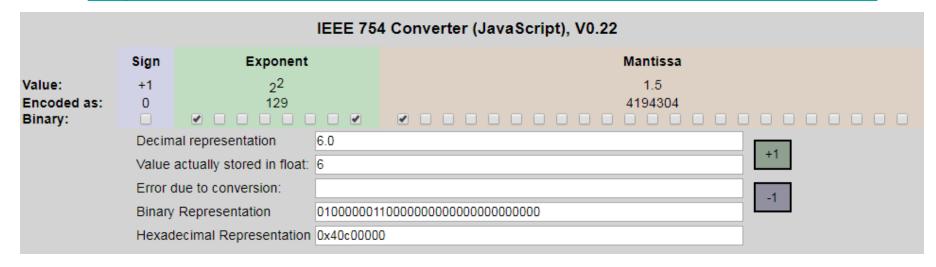


- - Sign: +
 - Exponent: 129 127 = +2
 - Mantissa: 100 0000...
- Decimal Value: $(+1.100\ 0000...)_2 \times 2^{+2} = 1.5 \times 2^{+2} = (+6.0)_{10}$

Useful Tool



- IEEE-754 Floating Point Converter
 - https://www.h-schmidt.net/FloatConverter/IEEE754.html



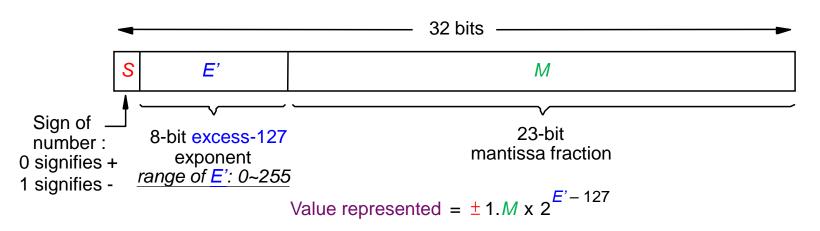
Class Exercise 2.6

Student ID: .	Date:
Name:	

- What is (-0.5)₁₀ in the IEEE single precision binary floating point format?
- Answer:

Special Values





- When exponent E' = 0 (all 0's) and mantissa M = 0:
 - The value 0 is represented.
- When exponent E' = 0 (all 0's) and mantissa $M \neq 0$:
 - Denormal values (i.e. very small values) are represented.
- When exponent E' = 255 (all 1's) and mantissa M = 0:
 - The value ∞ is presented.
- When exponent E' = 255 (all 1's) and mantissa $M \neq 0$:
 - Not a Number (NaN) (e.g. 0/0 or $\sqrt{-1}$) is presented.

Arithmetic on Floating-Point Number (1/2)

- When adding/subtracting floating-point numbers, their mantissas must be shifted with respect to each other.
 - E.g. adding $(2.9400)_{10} \times 10^2$ to $(4.3100)_{10} \times 10^4$
 - We rewrite $(2.9400)_{10} \times 10^2$ as $(0.0294)_{10} \times 10^4$
 - Then perform addition of the mantissas to get 4.3394×10^4 .

Add/Subtract Rule

- 1) Choose the number with the smaller exponent and shift its mantissa right a number of steps equal to the difference in exponents.
- 2) Set the exponent of the result equal to the larger exponent.
- Perform addition/subtraction on the mantissas and determine the sign of the result.
- 4) Normalize the resulting value, if necessary.

Arithmetic on Floating-Point Number (2/2)

- Multiplication and division are somewhat easier than addition and subtraction.
 - No alignment of mantissas is needed.
- Multiply Rule
 - 1) Add the exponents and subtract 127 to maintain the excess-127 representation.
 - 2) Multiply the mantissas and determine the sign of the result.
 - 3) Normalize the resulting value, if necessary.
- Divide Rule
 - 1) Subtract the exponents and add 127 to maintain the excess-127 representation.
 - 2) Divide the mantissas and determine the sign of the result.
 - 3) Normalize the resulting value, if necessary.

Outline



- Number Representation
 - Number Systems
 - Integers
 - Unsigned Integer
 - Signed Integer
 - Arithmetic Operations
 - Floating-Point Numbers
 - Unsigned Binary Fraction
 - Floating-Point Number Representation
 - Arithmetic Operations
- Character Representation
 - ASCII

Character Representation



- The most common encoding scheme for characters is ASCII (American Standard Code for Information Interchange).
- In ASCII encoding scheme, alphanumeric characters, operators, punctuation symbols, and control characters can be represented by 7-bit codes.
 - It is convenient to use an 8-bit byte to represent a character.
 - The code occupies the low-order 7 bits with the high-order bit as 0.

ASCII Table



Dec	Bin	Нех	Char	Dec	Bin	Нех	Char	Dec	Bin	Нех	Char	Dec	Bin	Нех	Char
0	0000 0000	00	[NUL]	32	0010 0000	20	space	64	0100 0000	40	9	96	0110 0000	60	``
1	0000 0001	01	[SOH]	33	0010 0001	21	•	65	0100 0001	41	A	97	0110 0001	61	a
2	0000 0010	02	[STX]	34	0010 0010	22	11	66	0100 0010	42	В	98	0110 0010	62	b
3	0000 0011	03	[ETX]	35	0010 0011	23	#	67	0100 0011	43	C	99	0110 0011	63	C
4	0000 0100	04	[EOT]	36	0010 0100	24	\$	68	0100 0100	44	D	100	0110 0100	64	d
5	0000 0101	05	[ENQ]	37	0010 0101	25	8	69	0100 0101	45	E	101	0110 0101	65	е
6	0000 0110	06	[ACK]	38	0010 0110	26	£	70	0100 0110	46	F	102	0110 0110	66	f
7	0000 0111	07	[BEL]	39	0010 0111	27	•	71	0100 0111	47	G	103	0110 0111	67	g
8	0000 1000	80	[BS]	40	0010 1000	28	(72	0100 1000	48	H	104	0110 1000	68	h
9	0000 1001	09	[TAB]	41	0010 1001	29)	73	0100 1001	49	I	105	0110 1001	69	i
10	0000 1010	0A	[LF]	42	0010 1010	2 A	*	74	0100 1010	4A	J	106	0110 1010	6A	j
11	0000 1011	0В	[VT]	43	0010 1011	2B	+	75	0100 1011	4 B	K	107	0110 1011	6B	k
12	0000 1100	0C	[FF]	44	0010 1100	2C	,	76	0100 1100	4C	L	108	0110 1100	6C	1
13	0000 1101	0D	[CR]	45	0010 1101	2D	_	77	0100 1101	4 D	M	109	0110 1101	6D	m
14	0000 1110	0E	[so]	46	0010 1110	2E	•	78	0100 1110	4E	N	110	0110 1110	6E	n
15	0000 1111	0F	[SI]	47	0010 1111	2 F	/	79	0100 1111	4F	0	111	0110 1111	6 F	0
16	0001 0000	10	[DLE]	48	0011 0000	30	0	80	0101 0000	50	P	112	0111 0000	70	p
17	0001 0001	11	[DC1]	49	0011 0001	31	1	81	0101 0001	51	Q	113	0111 0001	71	q
18	0001 0010	12	[DC2]	50	0011 0010	32	2	82	0101 0010	52	R	114	0111 0010	72	r
19	0001 0011	13	[DC3]	51	0011 0011	33	3	83	0101 0011	53	s	115	0111 0011	73	s
20	0001 0100	14	[DC4]	52	0011 0100	34	4	84	0101 0100	54	T	116	0111 0100	74	t
21	0001 0101	15	[NAK]	53	0011 0101	35	5	85	0101 0101	55	υ	117	0111 0101	75	u
22	0001 0110	16	[SYN]	54	0011 0110	36	6	86	0101 0110	56	v	118	0111 0110	76	v
23	0001 0111	17	[ETB]	55	0011 0111	37	7	87	0101 0111	57	W	119	0111 0111	77	w
24	0001 1000	18	[CAN]	56	0011 1000	38	8	88	0101 1000	58	X	120	0111 1000	78	x
25	0001 1001	19	[EM]	57	0011 1001	39	9	89	0101 1001	59	Y	121	0111 1001	79	У
26	0001 1010	1 A	[SUB]	58	0011 1010	3 A	:	90	0101 1010	5 A	${f z}$	122	0111 1010	7 A	z
27	0001 1011	1B	[ESC]	59	0011 1011	3B	;	91	0101 1011	5B]	123	0111 1011	7B	{
28	0001 1100	1C	[FS]	60	0011 1100	3C	<	92	0101 1100	5C	\	124	0111 1100	7C	1
29	0001 1101	1D	[GS]	61	0011 1101	3D	=	93	0101 1101	5D	1	125	0111 1101	7 D	}
30	0001 1110	1E	[RS]	62	0011 1110	3E	>	94	0101 1110	5E	^	126	0111 1110	7E	~
31	0001 1111	1 F	[ʊs]	63	0011 1111	3 F	?	95	0101 1111	5 F	_	127	0111 1111	7 F	[DEL]

Class Exercise 2.7



• Represent "Hello, CSCI2510" using ASCII code:

	Decimal	Binary
Н		
е		
1		
1		
0		
,		
С		
S		
С		
I		
2		
5		
1		
0		

Summary



- Number Representation
 - Number Systems
 - Integers
 - Unsigned Integer
 - Signed Integer
 - Arithmetic Operations
 - Floating-Point Numbers
 - Unsigned Binary Fraction
 - Floating-Point Number Representation
 - Arithmetic Operations
- Character Representation
 - ASCII