

การดำเนินการเกี่ยวกับตัวเลขและสายอักขระ

ดร.สายัณห์ สายยศ

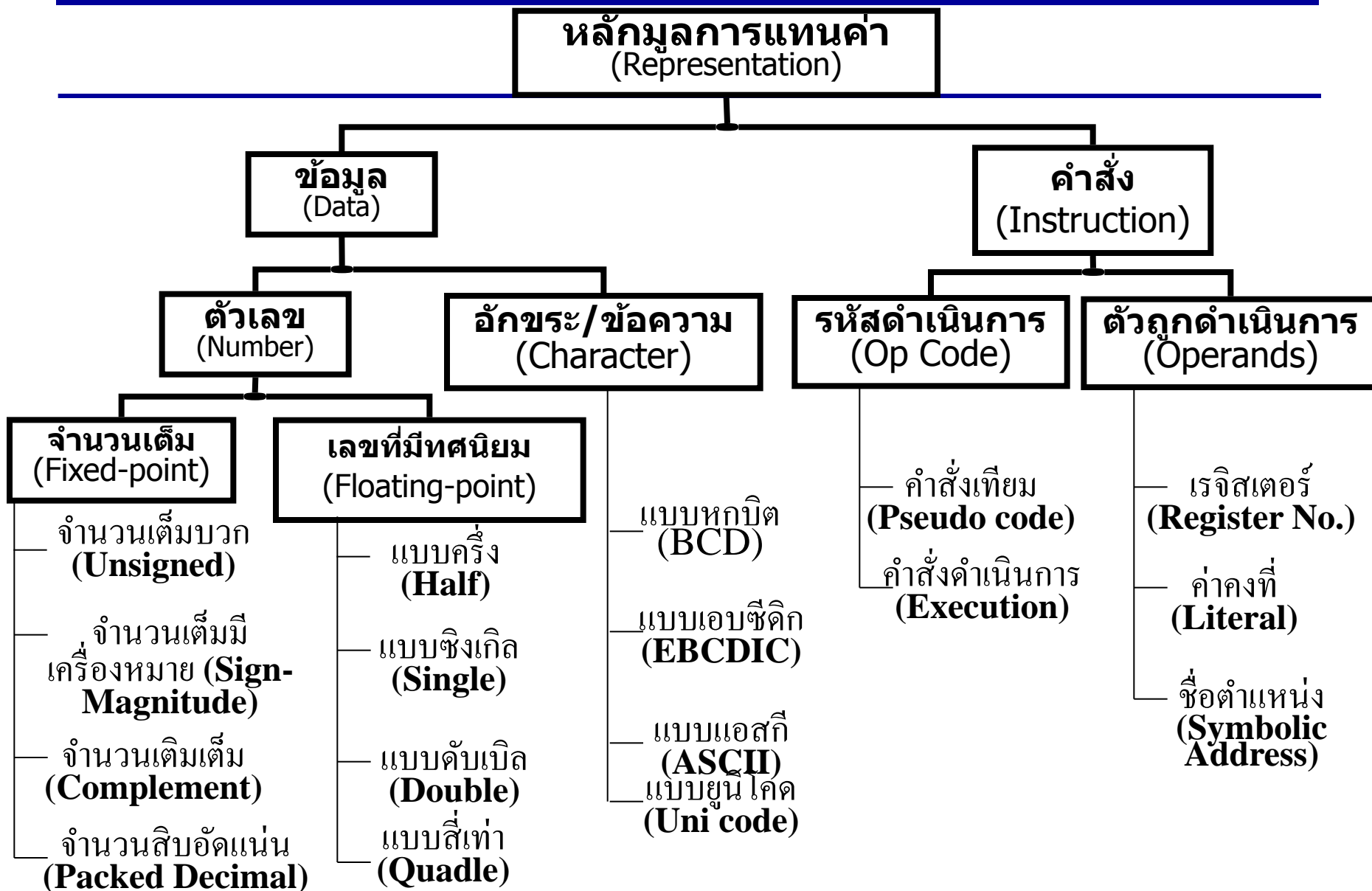
ภาควิชาวิทยาการคอมพิวเตอร์
คณะวิทยาศาสตร์ มหาวิทยาลัยขอนแก่น

วัตถุประสงค์การเรียนรู้

- เข้าใจการแปลงชนิดตัวแปรแบบตัวเลข
- การเรียกใช้ฟังก์ชันคณิตศาสตร์
- ฟังก์ชันการสุ่มตัวเลข
- การเข้าถึงค่าในสายอักขระ
- การดำเนินการเกี่ยวกับสายอักขระ
- การดำเนินการรูปแบบสายอักขระ
- ฟังก์ชันเกี่ยวกับสายอักขระ

หัวข้อที่น่าสนใจ

- บทนำ
- เข้าใจการแปลงชนิดตัวแปรแบบตัวเลข
- การเรียกใช้ฟังก์ชันคณิตศาสตร์
- ฟังก์ชันการสุมตัวเลข
- การเข้าถึงค่าในสายอักขระ
- การดำเนินการเกี่ยวกับสายอักขระ
- การดำเนินการรูปแบบสายอักขระ
- ฟังก์ชันเกี่ยวกับสายอักขระ



ระบบการแทนข้อมูล

- **Numeric Representation**

Fixed-Point Numbers (Integer)

Floating-Point Numbers (Real)

- **Character Representation**

BCD, ASCII, EBCDIC, UniCode

Numeric Representation

Fixed-point Numbers

- ➡ **Unsigned Binary**
- ➡ **Sign-Magnitude**
- ➡ **Complementation**
- ➡ **Packed Decimal**

Numeric Representation

เลขยกกำลังของสอง

$$2^0 = 1$$

$$2^1 = 2$$

$$2^2 = 4$$

$$2^3 = 8$$

$$2^4 = 16$$

$$2^5 = 32$$

$$2^6 = 64$$

$$2^7 = 128$$

$$2^8 = 256$$

$$2^9 = 512$$

$$2^{10} = 1,024$$

$$2^{11} = 2,048$$

$$2^{12} = 4,096$$

$$2^{13} = 8,192$$

$$2^7 = 128$$

$$2^8 = 256$$

$$2^9 = 512$$

$$2^{10} = K = 1,024$$

$$2^{15} = 32,768$$

$$2^{20} = M = 1,048,576$$

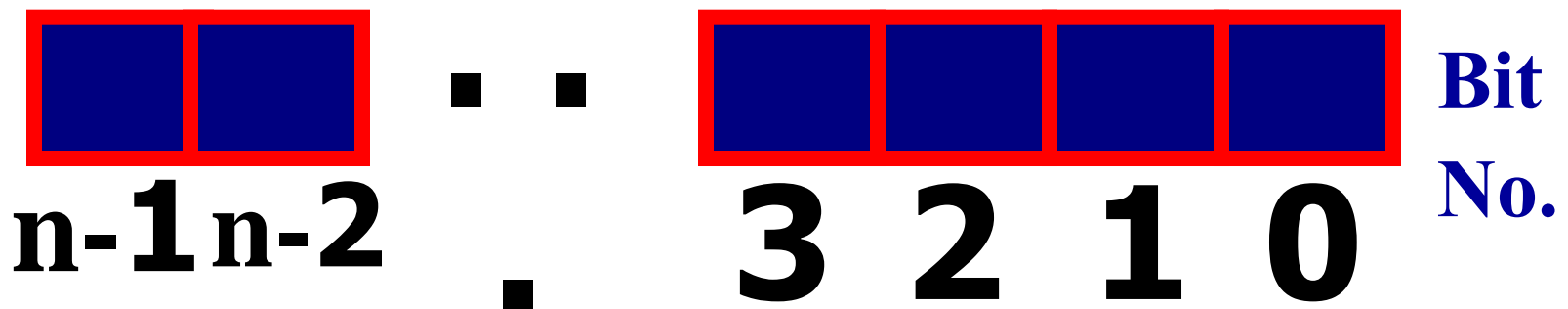
$$2^{30} = G = 1,073,741,824$$

Numeric Representation

Unsigned Binary

ใช้เลขฐานสองแทนค่า เลขจำนวนเต็ม โดยตรง

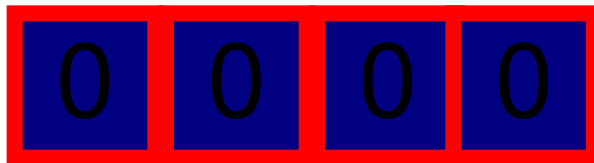
เมื่อ n = จำนวนบิตข้อมูลที่แทนค่า
ค่าที่แทนได้จะอยู่ในช่วง $0 - (2^n - 1)$



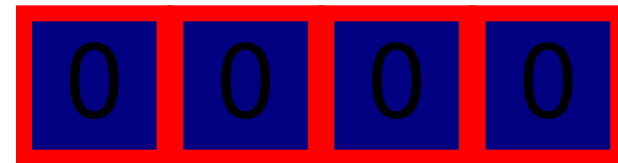
Numeric Representation

Unsigned Binary ตัวอย่างการแทนค่า

ถ้าใช้ 8 บิตในการแทนค่า (N=8)

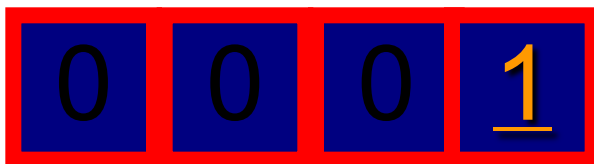


2^7 2^6 2^5 2^4

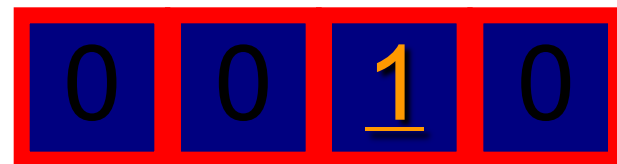


2^3 2^2 2^1 2^0

18 แทนด้วย 0001 0010



2^7 2^6 2^5 2^4



2^3 2^2 2^1 2^0

18

Numeric Representation

Unsigned Binary ตัวอย่างการแทนค่า

ถ้าใช้ 8 บิตในการแทนค่า ($N = 8$)

0 แทนด้วย 0000 0000

18 แทนด้วย 0001 0010

129 แทนด้วย 1000 0001
 2^7 2^0

255 แทนด้วย 1111 1111

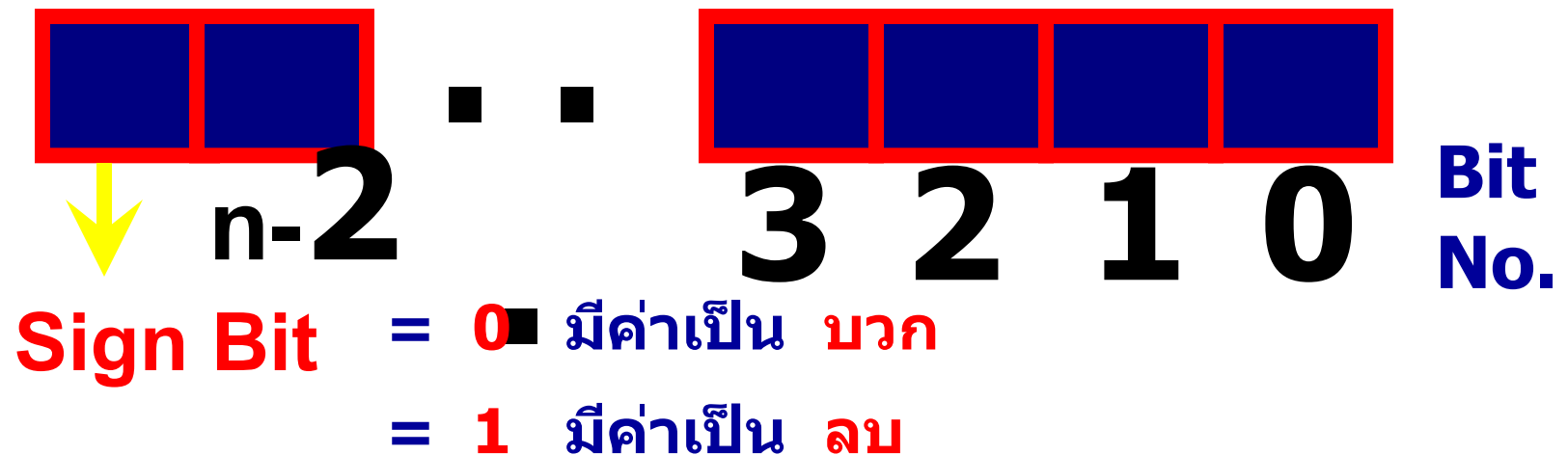
Numeric Representation

Sign-Magnitude

แทนค่าเลขจำนวนเต็มพร้อมเครื่องหมาย

ค่าที่แทนค่าได้อยู่ในช่วง

$$- (2^{n-1}-1) \text{ ถึง } + (2^{n-1}-1)$$

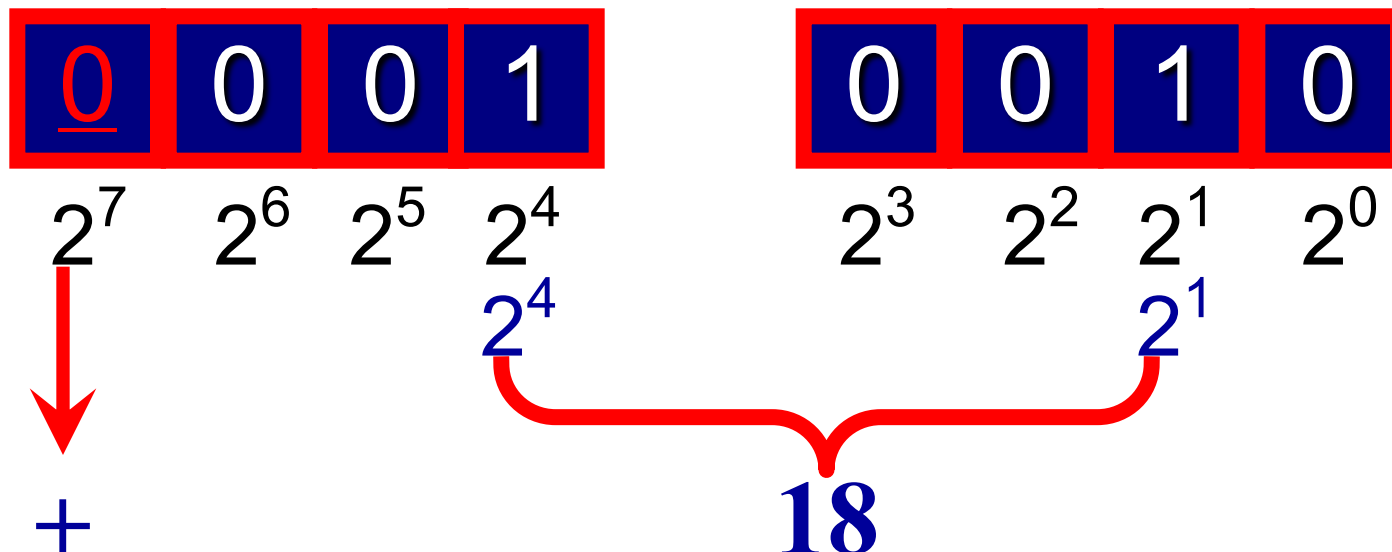


Numeric Representation

Sign-Magnitude

ถ้าใช้ 8 บิตในการแทนค่า ($N = 8$)

+18 แทนด้วย 0001 0010

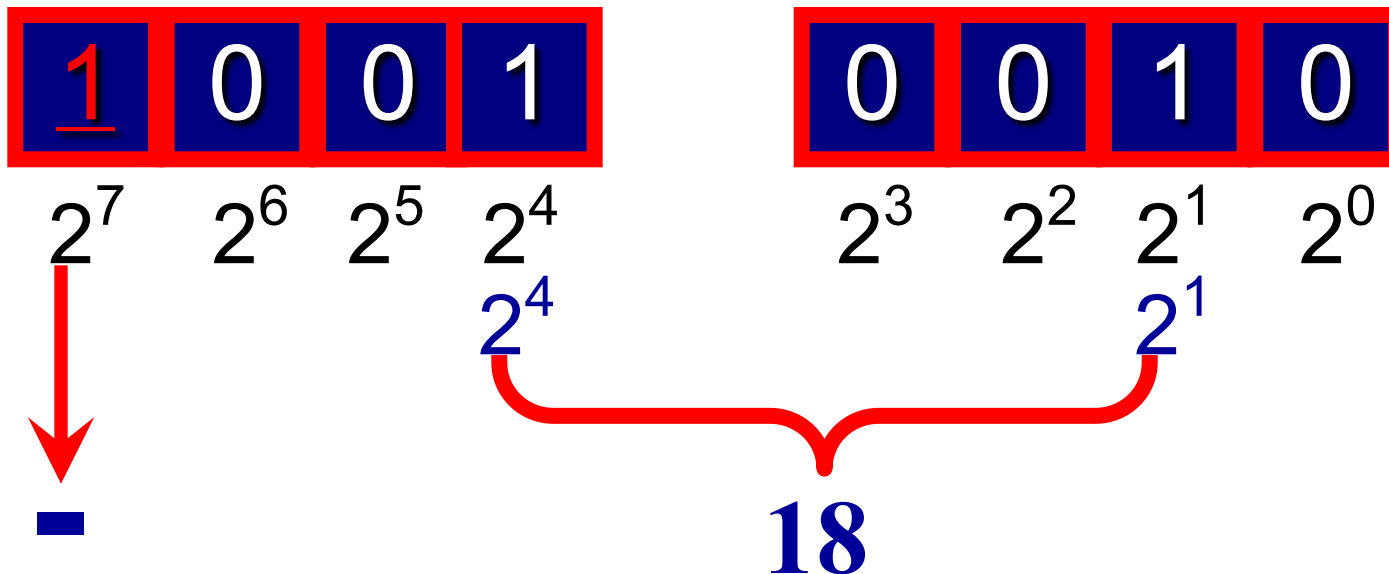


Numeric Representation

Sign-Magnitude ตัวอย่างการแทนค่า

ถ้าใช้ 8 บิตในการแทนค่า ($N = 8$)

-18 แทนด้วย 1001 0010



Numeric Representation

Sign-Magnitude ตัวอย่างการแทนค่า

ถ้าใช้ 8 บิตในการแทนค่า ($N = 8$)

+18 แทนด้วย 0001 0010

-18 แทนด้วย 1001 0010

+65 แทนด้วย 0100 0001

-65 แทนด้วย 1100 0001

+127 แทนด้วย 0111 1111

-127 แทนด้วย 1111 1111

Packed Decimal

ใช้ 4 บิตในการแทนค่าเลขฐานสิบ 1 หลัก

0000 = 0

0001 = 1

0010 = 2

0011 = 3

0100 = 4

0101 = 5

0110 = 6

0111 = 7

1000 = 8

1001 = 9

ใช้อีก 4 บิตแทนค่าเครื่องหมาย

1100 แทน บวก

1011 แทน ลบ

C

B

ตัวอย่างการแทนค่า

-567_{10} แทนด้วย

0101	0110	0111	1011
5	6	7	-

$+95,324_{10}$ แทนด้วย

1001	0101	0011	0010	0100	1100
9	5	3	2	4	+



Character Representation

ประเภทของตัวอักขระ

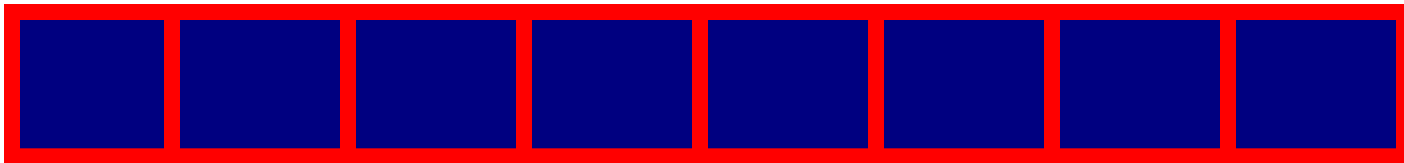
- ตัวอักษร (Alphabetic Characters)
- ตัวเลข (Numbers)
- ตัวคั่นวรรคตอน (Punctuations)
- สัญลักษณ์พิเศษ (Special Characters)
- ตัวอักขระควบคุม (Control Characters)
- ตัวอักขระกราฟิก (Graphic Characters)

Character Representation

ASCII

American Standard Code for Information Interchange

8-bits = Byte



ANSI : American National Standard Institute

Character Representation

ASCII

American **S**tandard **C**ode for **I**nformation **I**nterchange

ตัวอย่าง "A" แทนด้วย $41_{16} = 100\ 0001_2$

เมื่อใช้วิธี Odd Parity Checking

"A" จะแทนด้วย $C1_{16} = 1100\ 0001_2$

รหัสข้อมูล (Data Representation) (ต่อ)

ตารางรหัส ASCII แทนตัวอักษร

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	DLE	SP	0	@	P		P	๐	๑		ฐ	ภ	ะ	เ	อ
1	SOH	Dc1	1	1	A	Q	a	q	๒	๓	ก	ท	ม	๕	แ	ด
2	STX	Dc2	.	2	B	R	b	r	๔	๕	ข	ฌ	ย	า	โ	บ
3	ETX	Dc3		3	C	S	c	s	๖	๗	ฃ	ณ	ร	ำ	ใ	ค
4	EOT	Dc4	\$	4	D	T	d	t	๘	๙	ค	ด	ถ	๑	ไ	ล
5	ENQ	NAK	%	5	E	U	e	u	๙		ค	ด	ล	๑		ล
6	ACK	SYN	&	6	F	V	f	v	๙	—	ฅ	ถ	ภ	๑	๗	๖
7	BEL	ETB	.	7	G	W	g	w	๙	+	ง	ท	ว	๑	๙	๗
8	BS	CAN	(8	H	X	h	x	๙	┌	จ	ธ	ศ	.	.	ณ
9	HT	EM)	9	I	Y	i	y	๙	└	ฉ	น	ษ	๑	๑	ด
A	LF	SUB	.	:	J	Z	j	z	๙	└	ช	บ	ส	.	๑	
B	VT	ESC	+	:	K	[k	{	๙	└	ช	ป	ท		.	
C	FF	FS	.	<	L	\	l	l	๙	└	ฌ	ผ	พ		๑	
D	CR	GS	-	=	M]	m	}	๙	└	ญ	ผ	อ		.	
E	SO	RS	.	>	N	^	n	-	๙	└	ญ	พ	ฮ			
F	SI	US	/	?	O	-	o	M	๙	└	ญ	ฟ	ย	฿	●	

Character Representation

ASCII

American Standard Code for Information Interchange

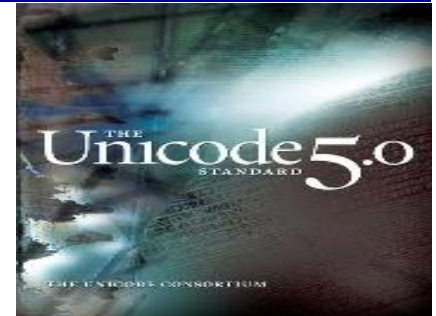
- ปรับมาตรฐาน 7 Bits => ISO/IEC 646
- **ISO/IEC 10646-1991**
- 8 Bits => ISO/IEC 8859-1-1998

Character Representation

การเข้ารหัสยูนิโคด Unicode

The Unicode Consortium

<http://www.unicode.org>



แทนค่าตัวอักขระได้หลากหลาย 30 กว่าภาษาทั่วโลก โดยเป็นตัวอักขระ พยัญชนะ สระ ตัวสะกดต่างๆ ทั้งที่ใช้ในการเขียนประโยคข้อความทั่วไป เครื่องหมายหรือสัญลักษณ์ต่างๆ ที่ใช้คำนวณวรรคตอน เครื่องหมายทางคณิตศาสตร์ ตัวโน้ตดนตรี ฯลฯ

เข้าใจการแปลงชนิดตัวแปรแบบตัวเลข

Python supports four different numerical types:

- **int (signed integers)**: They are often called just integers or ints, are positive or negative whole numbers with no decimal point.
- **long (long integers)**: Also called longs, they are integers of unlimited size, written like integers and followed by an uppercase or lowercase L.
- **float (floating point real values)** : Also called floats, they represent real numbers and are written with a decimal point dividing the integer and fractional parts. Floats may also be in scientific notation, with E or e indicating the power of 10 ($2.5e2 = 2.5 \times 10^2 = 250$).
- **complex (complex numbers)** : are of the form $a + bJ$, where a and b are floats and J (or j) represents the square root of -1 (which is an imaginary number). The real part of the number is a, and the imaginary part is b. Complex numbers are not used much in Python programming.

เข้าใจการแปลงชนิดตัวแปรแบบตัวเลข

Number Type Conversion

Python converts numbers internally in an expression containing mixed types to a common type for evaluation. But sometimes, you need to coerce a number explicitly from one type to another to satisfy the requirements of an operator or function parameter.

- Type **int(x)** to convert x to a plain integer.
- Type **long(x)** to convert x to a long integer.
- Type **float(x)** to convert x to a floating-point number.
- Type **complex(x)** to convert x to a complex number with real part x and imaginary part zero.

ฟังก์ชันการสุ่มตัวเลข

Function	Description
<u>choice(seq)</u>	A random item from a list, tuple, or string.
<u>randrange ([start,] stop [,step])</u>	A randomly selected element from range(start, stop, step)
<u>random()</u>	A random float r, such that 0 is less than or equal to r and r is less than 1

การเข้าถึงค่าในสายอักขระ

```
var1 = 'Hello World!'
```

```
var2 = "Python Programming"
```

```
print "var1[0]: ", var1[0]
```

```
print "var2[1:5]: ", var2[1:5]
```

การดำเนินการเกี่ยวกับสายอักขระ

Operator	Description	Example
+	Concatenation - Adds values on either side of the operator	a + b will give HelloPython
*	Repetition - Creates new strings, concatenating multiple copies of the same string	a*2 will give -HelloHello
[]	Slice - Gives the character from the given index	a[1] will give e
[:]	Range Slice - Gives the characters from the given range	a[1:4] will give ell
in	Membership - Returns true if a character exists in the given string	H in a will give 1
not in	Membership - Returns true if a character does not exist in the given string	M not in a will give 1

การดำเนินการรูปแบบสายอักขระ

One of Python's coolest features is the string format operator `%`. This operator is unique to strings and makes up for the pack of having functions from C's `printf()` family. Following is a simple example:

```
#!/usr/bin/python  
  
print "My name is %s and weight is %d kg!" % ('Zara', 21)
```

When the above code is executed, it produces the following result:

```
My name is Zara and weight is 21 kg!
```

Here is the list of complete set of symbols which can be used along with `%`:

การดำเนินการรูปแบบสายอักขระ

Format Symbol	Conversion
%c	character
%s	string conversion via str() prior to formatting
%i	signed decimal integer
%d	signed decimal integer
%u	unsigned decimal integer

ฟังก์ชันเกี่ยวกับสายอักขระ

Sr. No.	Methods with Description
1	<u>capitalize()</u> Capitalizes first letter of string.
2	<u>center(width, fillchar)</u> Returns a space-padded string with the original string centered to a total of width columns.
3	<u>count(str, beg= 0,end=len(string))</u> Counts how many times str occurs in string or in a substring of string if starting index beg and ending index end are given.
4	<u>decode(encoding='UTF-8',errors='strict')</u> Decodes the string using the codec registered for encoding. encoding defaults to the default string encoding.
5	<u>encode(encoding='UTF-8',errors='strict')</u> Returns encoded string version of string; on error, default is to raise a ValueError unless errors is given with 'ignore' or 'replace'.

ฟังก์ชันเกี่ยวกับสายอักขระ

8	<u><code>find(str, beg=0 end=len(string))</code></u> Determine if str occurs in string or in a substring of string if starting index beg and ending index end are given returns index if found and -1 otherwise.
9	<u><code>index(str, beg=0, end=len(string))</code></u> Same as find(), but raises an exception if str not found.

ฟังก์ชันเกี่ยวกับสายอักขระ

10	<u>isalnum()</u> Returns true if string has at least 1 character and all characters are alphanumeric and false otherwise.
11	<u>isalpha()</u> Returns true if string has at least 1 character and all characters are alphabetic and false otherwise.
12	<u>isdigit()</u> Returns true if string contains only digits and false otherwise.
13	<u>islower()</u> Returns true if string has at least 1 cased character and all cased characters are in lowercase and false otherwise.
14	<u>isnumeric()</u> Returns true if a unicode string contains only numeric characters and false otherwise.
15	<u>isspace()</u> Returns true if string contains only whitespace characters and false otherwise.

ฟังก์ชันเกี่ยวกับสายอักขระ

31	<u><code>split(str="", num=string.count(str))</code></u> Splits string according to delimiter str (space if not provided) and returns list of substrings; split into at most num substrings if given.
32	<u><code>splitlines(num=string.count("\n"))</code></u> Splits string at all (or num) NEWLINEs and returns a list of each line with NEWLINEs removed.

สรุป

- บทนำ
- เข้าใจการแปลงชนิดตัวแปรแบบตัวเลข
- การเรียกใช้ฟังก์ชันคณิตศาสตร์
- ฟังก์ชันการสุมตัวเลข
- การเข้าถึงค่าในสายอักขระ
- การดำเนินการเกี่ยวกับสายอักขระ
- การดำเนินการรูปแบบสายอักขระ
- ฟังก์ชันเกี่ยวกับสายอักขระ

