爬蟲基礎入門-Part 2

Python 安裝

Windows

https://www.python.org/

Ubuntu

sudo apt-get install python3.6

Mac

https://docs.python-guide.org/en/latest/starting/install/osx/

Python 版本

Python 2 vs. Python 3?

為什麼用Python 2

　　許多函式庫尚未完成 Python 3 的移植

　　Python 2 最後一個開發版本將維持在 2.7.10，並且不會再有新版（不再加新功能），除非有重大安全才會釋
　　出安全更新版。

為什麼用Python 3

　　更強大更方便的語法

　　更強大更方便的函式庫

　　Python 3 是目前官方持續開發（加新功能）的版本。
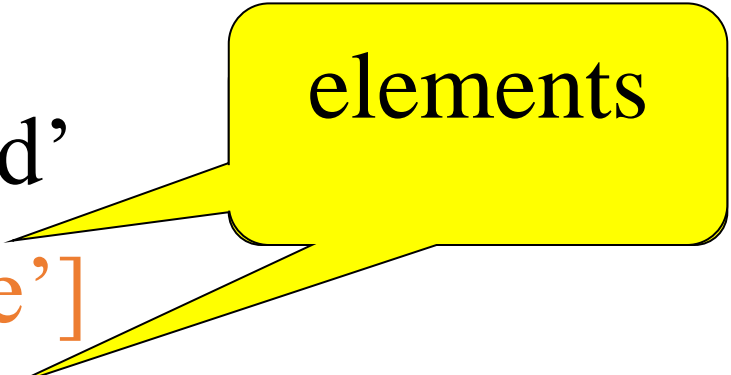
基本輸入輸出

基本輸出：print()

```
1   school="CKSH"
2   year=2013
3
4   #可以用 , 分隔多個變數一起輸出
5   print( "Hello", school, year, "!" )
6   #預設 分隔,結束 分別為 sep=' ', end='\n'
7   #你可以修改成你想要的
8   print( "Hello", school, year, "!", sep='_', end='\nend\n' )
9
10  for_pause = input("press enter to exit")
```

小知識：python的註解

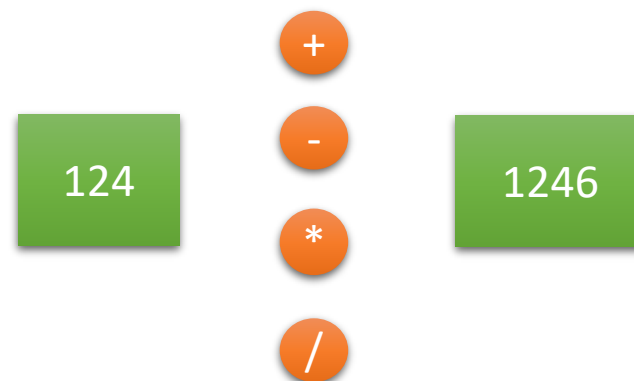- 單行註解：**#**之後到換行為止都是註解（所以最多註解單行）
- 多行註解：利用多行字串 """ ... """

```
1  num = 1246
2  #我是註解喔^^
3  print( 1246*3 )  #我也是註解:3
4  """
5  我是多行註解 ^.<
6  第二行
7  第三行
8  """
```

Built-in DataTypes

- Number
- String
- List
- Dictionary
- Set

7, 11.0

'Hello World'

[7, 11, 'store']

{'name':'marr', 'home':'taipei'}

set([7, 11, 'store'])

elements

int

整數

+

-

124   *   1246

/

新增整數變數、四則運算
Elementary arithmetic

```
1  num1 = 2
2  num2 = 5
3  print( num1.__class__ )  #只是確認看看是不是 "int"
4  print( num2.__class__ )  #同上
5
6  print( "num1+num2=", num1+num2 )
7  print( "num1-num2=", num1-num2 )
8  print( "num1*num2=", num1*num2 )
9  print( "num1/num2=", num1/num2 )    #注意這裡不是無條件捨去
10 print( "num1//num2=", num1//num2 )  #無條件捨去
11 print( "num1%num2=", num1%num2 )
12 print( "num1**num2=", num1**num2 )
13
14 for_pause = input('Press enter to exit')
```

指定運算子 =

```python
#可以一次多個指定
n, name, f = 5, "rock", 1.02**10
print( n, name, f )

for_pause = input("press enter to exit")
```

string

字串

| 'H' | 'a' | 't' | 's' | 'u' | 'n' | 'e' | 'M' | 'l' | 'k' | 'u' |

新增字串變數、加乘運算
string operation

```python
1   #單引號 或 雙引號都是string
2   #可以視字串值是否包含單雙引號決定要用哪種
3   name='Christina'
4   item="Steins;Gate"
5   action="is opened."
6
7   #加法就是字串串接
8   sentence = name + "'s " + item + ' ' + action
9   print( sentence )
10
11  #乘法=做很多次加法
12  print( (name+'! ')*3 )
13
14  for_pause = input("press enter to exit")
```

String
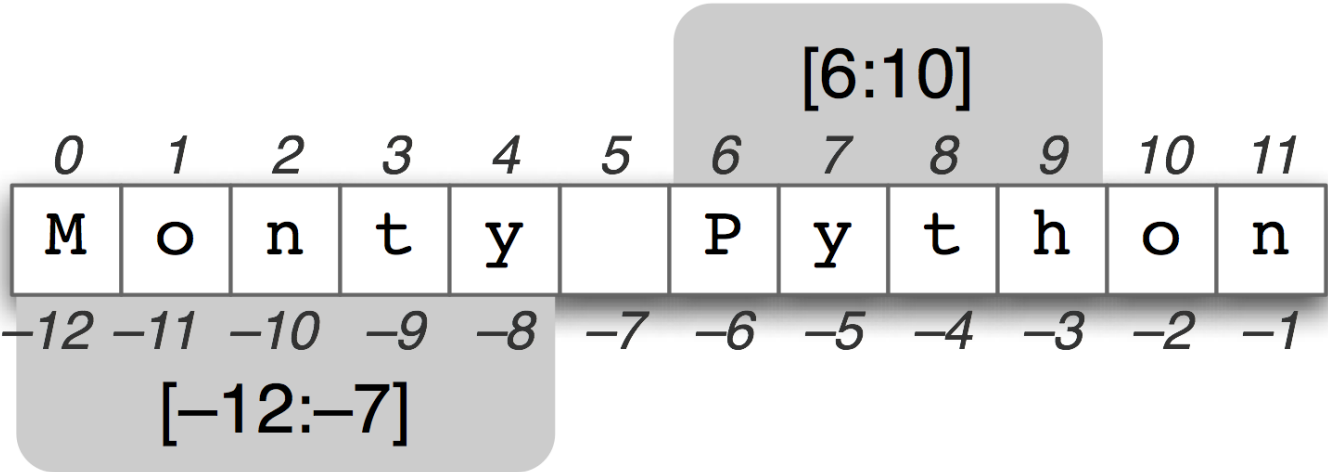
Control Flow

程式流程控制：if、while

- No explicit begin or end
- No curly braces

```
if age > 18:
    print("You are old enough.")
else:
    print("You are too young.")
```

if Statements

- ```python
  x = int(raw_input("Please enter an integer: "))
  if x < 0:
      x = 0
      print('Negative changed to zero')
  elif x == 0:
      print('Zero')
  elif x == 1:
      print('Single')
  else:
      print('More')
  ```

For Statements

- words = ['cat', 'window', 'defenestrate']
- for w in words:
      print(w, len(w))
  ➜ cat 3 window 6 defenestrate 12
- range(5, 10) ➜ [5, 6, 7, 8, 9]
- range(0, 10, 3) ➜ [0, 3, 6, 9]
- xrange vs. range
- for i in xrange(10):
      print (i)
  ➜ 0 1 2 3 4 5 6 7 8 9

while Statements

- ```
  x = 0
  while x < 10:
      print x,
      x += 1
  ```

  ➔ 0 1 2 3 4 5 6 7 8 9

Data Structure

資料結構：List、Dictionary、Set

List

- `squares = [1, 4, 9, 16, 25]`
- `len(squares)` ➜ `5`
- `squares[-1]` ➜ `25`
- `squares[-3:]` ➜ `[9, 16, 25] # slicing returns a new list`
- `squares[:]` ➜ `[1, 4, 9, 16, 25] # slicing returns a new list`
- `squares + [36, 49]` ➜ `[1, 4, 9, 16, 25, 36, 49]`
- `squares.append(36)` ➜ `[1, 4, 9, 16, 25, 36]`
- `squares[2:3] = []` ➜ `[1, 4, 25]`
- `squares[:] = []`
- `x = [['a', 'b', 'c'], [1, 2, 3]]`

Dictionary

- `tel = {'jack': 4098, 'sape': 4139}`
- `tel['guido'] = 4127`
  ➜ `tel = {'jack': 4098, 'sape': 4139, 'guido': 4127}`
- `del tel['sape']`
  ➜ `tel = {'jack': 4098, 'guido': 4127}`
- `tel.keys()`
  ➜ `['guido', 'irv', 'jack']`
- `'guido' in tel`
  ➜ `True`

Set

- basket = ['apple', 'orange', 'apple', 'pear', 'orange', 'banana']
- fruit = set(basket)   # create a set without duplicates
- 'orange' in fruit ➔ True
- a = set('abracadabra')
  ➔ set(['a', 'r', 'b', 'c', 'd'])
- a = set('abc')   b = set('ade')
- a – b ➔ set(['c', 'b'])
- a | b ➔ set(['a', 'c', 'b', 'e', 'd'])
- a & b ➔ set(['a'])
- a ^ b ➔ set(['c', 'b', 'e', 'd'])

function

函式（函數）

Function Example

```
def myFunc(a, b, c):     內部變數要讓外部也知道的參數
    global r
    x = a + 1
    y = b + 2
    z = C + 3                內部變數
    r = x**2+y**2+z**2
        return x, y, z

Print(r)
```