

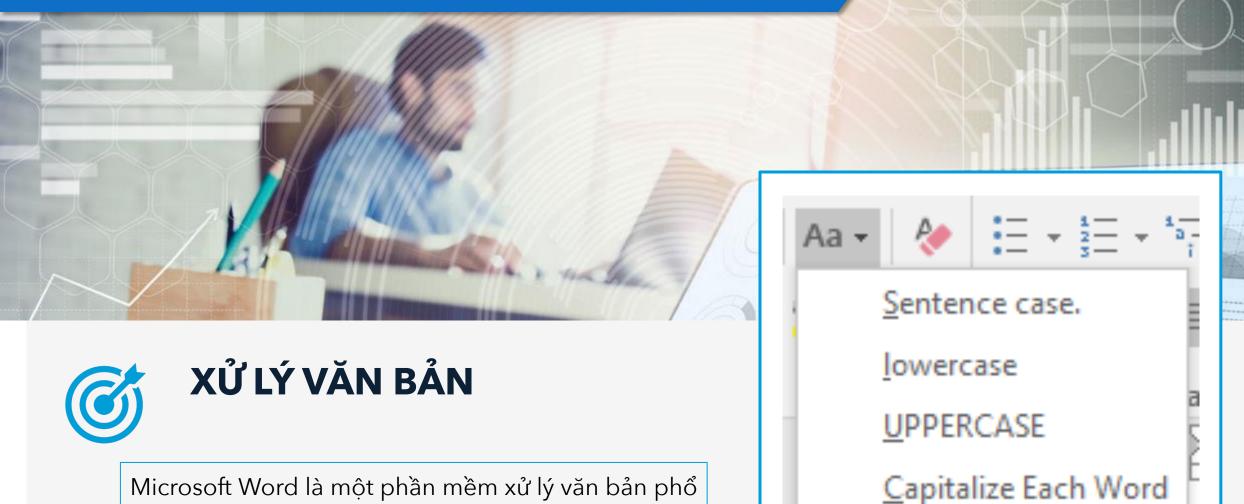
NHẬP MÔN KHOA HỌC MÁY TÍNH

- Phần 1: Khoa học máy tính trong thời đại số
- Phần 2: Lập trình máy tính giải quyết bài toán
 - Giới thiệu ngôn ngữ lập trình Python
 - Một số thành phần cơ bản của ngôn ngữ lập trình Python
 - Câu lệnh rẽ nhánh
 - Câu lệnh lặp
 - Hàm
- Phần 3: Giải quyết một số bài toán thực tế trên máy tính



KIỂU DỮ LIỆU XÂU KÍ TỰ

KHOA CÔNG NGHỆ THÔNG TIN TRƯỜNG ĐẠI HỌC SƯ PHẠM HÀ NỘI



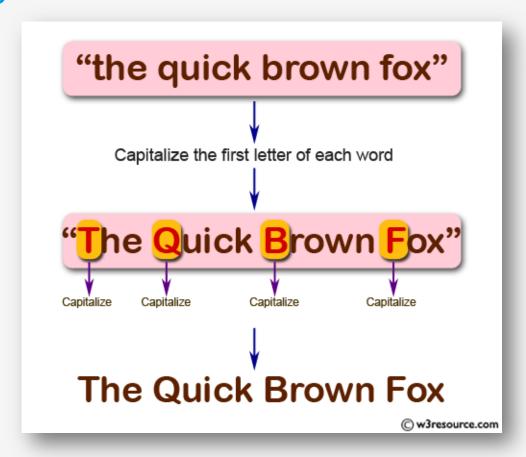
tOGGLE cASE

Microsoft Word là một phần mềm xử lý văn bản phổ biến với chức năng chính là soạn thảo và định dạng văn bản như: thay đổi font chữ, ...



NỘI DUNG

- Cấu trúc dữ liệu xâu kí tự, string
 - Định nghĩa và khai báo
 - Truy cập các kí tự trong xâu
- Các phép toán
- Các thao tác với xâu
- Các hàm thư viện xử lý xâu



KIỂU XÂU KÍ TỰ STRING



KIỂU XÂU KÍ TỰ STRING

- Là một kiểu dữ liệu cơ bản trong Python, dùng để lưu trữ một danh sách các kí tự liên tiếp nhau.
- Các ký tự được viết trong cặp dấu nháy đơn (') hoặc cặp dấu nháy kép (") hoặc cặp ba dấu nháy đơn ("') hoặc cặp ba dấu nháy kép (""")
- Xâu có thể rỗng (gồm 0 ký tự)
- Python không có kiểu kí tự, thay vào đó kí tự được coi là xâu có độ dài 1, chỉ bao gồm một kí tự

```
str1 = "Hello"
str2 = 'there'
str = str1 + str2
print(str)
str3 = ""
print(str3)
print(str4)
```



XÂU KÍ TỰ STRING

- Xâu chứa các chữ số khác với một số
 - "123" != 123
- Có thể chuyển một xâu kí tự số thành một số tương ứng và ngược lại.
- Tất cả dữ liệu đọc vào từ bàn phím đều được lưu dưới dạng xâu kí tự
 - Cần chuyển đổi kiểu phù hợp trước khi sử dụng

```
>>>str = '123'
>>>str = str + 1
Traceback (most recent call last):
File "<input>", line 1, in <module>
TypeError: can only concatenate str
(not "int") to str
>>>x = int(str) + 1
>>>print(x)
124
>>>y = float("1234.5")
>>>print(y)
1234.5
```



MỘT SỐ KÍ TỰ ĐẶC BIỆT

- Ngoài các kí tự thông thường (1 kí tự), python còn cung cấp các kí tự đặc biệt (gọi là escape sequence)
- Kí tự đặc biệt là dãy các kí tự bao gồm dấu gạch phải (\) và theo sau là một kí tự hoặc dãy số.
- Hãy cho biết kết quả hiển thị đoạn code dưới đây:

Kí tự	Nội dung
\n	Xuống dòng
\r	Về đầu dòng (Enter)
\t	Tab
\v	Tab dọc
/b	Lùi trái (Backspace)
\f	Hết form
\'	Nháy đơn '
\"	Nháy kép "
\\	Dấu sổ ngược \

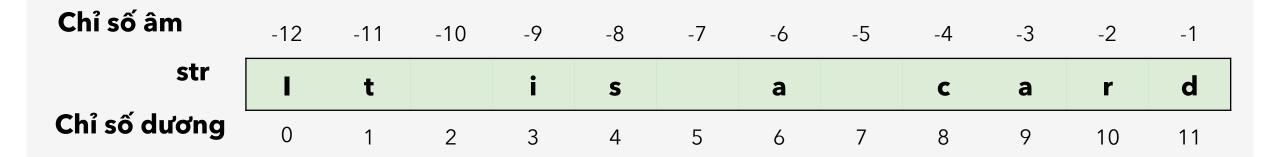
```
str = "Me di lam\n\tTu sang som"
print(str)
```

```
str = "Me bao: \"Me di lam nhe\""
print(str)
```





- Các kí tự trong xâu được tham chiếu đến bằng cách sử dụng chỉ số (index) đặt bên trong dấu ngoặc vuông([]).
- Chỉ số dương của xâu kí tự được bắt đầu từ 0, tăng dần từ trái sang phải
- Chỉ số âm bắt đầu từ -1, giảm dần từ phải sang trái





TRUY CẬP CÁC PHẦN TỬ

- Nếu chỉ số nằm ngoài phạm vi cho phép, lỗi sẽ xảy ra
- Nếu xâu có độ dài n, chỉ số nằm trong khoảng [0, n)
- Lấy độ dài của xâu kí tự bằng làm len(<xau ki tu>)

```
>>> st = "Day la cai ban"
>>> st[1]
555
>>> st[16]
Traceback (most recent call last):
  File "<pyshell>", line 1, in <module>
IndexError: string index out of range
>>> st[-3]
555
>>> st[-20]
Traceback (most recent call last):
  File "<pyshell>", line 1, in <module>
IndexError: string index out of range
```



DUYỆT XÂU KÍ TỰ

- Có 2 cách để duyệt xâu:
 - Duyệt trực tiếp qua từng kí tự với vòng lặp for
 - Duyệt dựa trên các chỉ số với vòng lặp for hoặc while
 - Cần chú ý độ dài xâu để không truy cập vào các vị trí nằm ngoài phạm vi cho phép.
 - Nên sử dụng hàm len()

```
st = "Ten toi la Hoa"

for ch in st:
    print(ch, end=' ')

for i in range(len(st)):
    print(st[i], end=' ')
```



VÍ DỤ 7-1

Nhập vào một chuỗi kí tự là họ tên của một người, hãy kiếm tra xem trong họ tên vừa nhập có chứa kí tự 'H' hoặc 'h' không?

```
hoten = input("Cho biết họ tên: ")
   print(f"Ho tên ban là: {hoten}")
   check = False
   for i in range(len(hoten)):
       if (hoten[i] == 'H') or (hoten[i] == 'h'):
           check = True
           break
   if (check == True):
       print("Trong tên của bạn có kí tự H hoặc h")
10
   else:
11
       print("Tên của bạn không có kí tự H hoặc h")
12
```

Cho biết họ tên: Nguyễn Nhật Minh Quang Họ tên bạn là: Nguyễn Nhật Minh Quang Trong tên của bạn có kí tự H hoặc h





- Xâu kí tự là bất biến
 - Không thể thêm hay bớt kí tự.
 - Không thể thay đổi kí tự của xâu
- Khi thực hiện các phép toán trên xâu, ví dụ phép nối xâu '+', một xâu mới sẽ được tạo ra để gán vào biến hiện tại. Ví dụ
 - s = "Hello"
 - s = s + "Everybody"

s = "nham Dan"

Lệnh	Kết quả
s[0] = "N"	Ð/S
s = "Nham Ty"	Ð/S
s = "2022 la" + s	Ð/S
s = "123" + 1	Ð/S
print(s[20])	Ð/S



VÍ DŲ 7.2

 Cho một số nguyên dương N (N<=10⁸), số N được gọi là đối xứng nếu viết từ trái sang phải hoặc từ phải sang trái thì giá trị không thay đổi

- Ví dụ:
 - Số 12321, 363, 5665 là các số đối xứng
 - Số 12333, 5661, 366 không phải là số đối xứng.

STT	N	dv	M
1	5665	#	0
2	566	5	5
3	56	6	56
4	5	6	566
5	0	5	5665





```
s = input("Hay nhap vao 1 so:")
invs = "" # so viet nguoc lai
for i in range(???):
     invs = invs + s[i]
print(f"So viet nguoc lai: {invs}")
if s == invs:
     print(f"So doi xung")
else:
     print(f"So KHONG doi xung")
```

```
Hay nhap vao 1 so:245
542
So KHONG doi xung
```

```
Hay nhap vao 1 xau:12321
12321
So doi xung
```





Bài 7.1: Hãy viết chương trình thực hiện công việc sau:

- a. Cho phép nhập vào một xâu gồm các chữ số
- b. Đếm và hiển thị số lần xuất hiện của chữ số '0' trong xâu vừa nhập
- c. *Thống kê tất cả các chữ số và số lần xuất hiện của chữ số đó trong xâu.

N = "1000101010010110"
So lan xuat hien cua 0: 9
Cac chu so va so lan xuat hien:
0: 9

N = "2312312321232"
So lan xuat hien cua 0: 9
Cac chu so va so lan xuat hien:
1: 3

2: 6

1: 6

CÁC PHÉP TOÁN



CÁC PHÉP TOÁN CƠ BẢN

Toán tử	Cú pháp	Mô tả
+	str1 + str2	Nối 2 xâu kí tự <i>str1</i> và <i>str2</i> để tạo thành một xâu mới chứa cả hai xâu
*	str * <n></n>	Tạo một xâu kí tự bằng cách nối n bản sao của xâu <i>str</i> với nhau
[:]	str[start:stop:step]	Trả về một xâu con chứa các kí tự trong str bắt đầu từ <start> đến <stop>-1. Dấu ":" bắt buộc phải có, các thành phần start, stop và step có thể không có. Nếu không có, mặc định start=0; stop=len(str); step=1.</stop></start>
in	x in str	Trả về True nếu xâu x có xuất hiện trong xâu <i>str</i>
not in	x not in str	Trả về True nếu xâu x không có trong xâu <i>str</i>







$$s1 = s[3:8]$$





$$s1 = s[3: 100]$$





$$s1 = s[:]$$





$$s2 = "ban"$$

STT	Phép toán	Kết quả
1	s = s1 + s2	???
2	s = s1*3	???
3	s = s1[1:2]	???
4	s = s1[1:]	???
5	s = s1[-1:-3]	???
6	s = s1[-1:-3:-1]	???
7	"Un" in s	





- Cho một số nguyên dương N (N<=10⁸), số N được gọi là đối xứng nếu viết từ trái sang phải hoặc từ phải sang trái thì giá trị không thay đổi
- Ví dụ:
 - Số 12321, 363, 5665 là các số đối xứng
 - Số 12333, 5661, 366 không phải là số đối xứng.

STT	N	dv	M
1	5665	#	0
2	566	5	5
3	56	6	56
4	5	6	566
5	0	5	5665



VÍ DỤ 7.2 – TỐI ƯU CHƯƠNG TRÌNH

```
s = input("Hay nhap vao 1 so:")
invs = "" # so viet nguoc lai
for i in range(len(s)-1, -1, -1):
      invs = invs + s[i]
print(f"So viet nguoc lai: {invs}")
if s == invs:
      print(f"So doi xung")
else:
      print(f"So KHONG doi xung")
```

```
s = input("Hay nhap vao 1 so:")
# tim so viet nguoc lại
invs = s[::-1]
print(f"So viet nguoc lai: {invs}")
if s == invs:
      print(f"So doi xung")
else:
      print(f"So KHONG doi xung")
```



CÁC PHÉP TOÁN QUAN HỆ

Toán tử	Cú pháp	Mô tả
==	str1 == str2	Trả về <i>True</i> nếu hai xâu <i>str1</i> và <i>str2</i> giống nhau
!=	str1!= str2	Trả về <i>True</i> nếu hai xâu <i>str1</i> và <i>str2</i> không giống nhau
>	str1 > str2	Trả về <i>True</i> nếu xâu <i>str1</i> lớn hơn <i>str2</i>
>=	str1 >= str2	Trả về <i>True</i> nếu xâu <i>str1</i> lớn hơn hoặc bằng <i>str2</i>
<	str1 < str2	Trả về <i>True</i> nếu xâu <i>str1</i> nhỏ hơn <i>str2</i>
<=	str1 <= str2	Trả về <i>True</i> nếu xâu <i>str1</i> nhỏ hơn hoặc bằng <i>str2</i>

Python so sánh hai xâu dựa trên mã ASCII của kí tự, bắt đầu từ trái sang phải. Và mọi xâu đều lớn hơn xâu rỗng





Nguyên lý so sánh hai xâu:

$$A = a_0 a_1 \dots a_m \text{ và } B = b_0 b_1 \dots b_n$$

- So sánh từ đầu theo từng cặp ký tự (a_i, b_i) , i = 1,2,3,...
 - Nếu gặp $a_i < b_i$, kết luận ngay a < b, stop
 - Nếu gặp $a_i > b_i$, kết luận ngay a > b, stop
- Nếu không thấy cặp ký tự tương ứng khác nhau, xâu ngắn hơn được coi là nhỏ hơn

```
"computer" > "calculator"

"COMPUTER" < "calculator"

"professional" < "professor"

"10" < "9"

"with" < "without"
```

Tốc độ phép so sánh phụ thuộc vào vị trí i đầu tiên mà $a_i \neq b_i$





s1 = "Hanh"

s2 = "Hanna"

STT	Phép toán	Kết quả
1	s1 == s2	???
2	s1 != s2	???
3	s1 > s2	???
4	s1 >= s2	???
5	s1 < s2	???
6	s1 <= s2	???
7	s1 >= "Hanhit"	???





Bài 7.2: Hãy viết chương trình cho phép thực hiện các công việc sau:

- Nhập vào tên của ba thành phố trong nước
- Sắp xếp các thành phố này theo thứ tự ABC
- Sắp xếp các thành phố này theo thứ tự ABC không kể viết hoa hay thường

Tp1: Hai phong

Tp2: Quang Ninh

Tp3: Ha noi

Sau khi sap xep:

Ha noi

Hai phong

Quang Ninh

CÁC HÀM THƯ VIỆN



CÁC HÀM CƠ BẢN

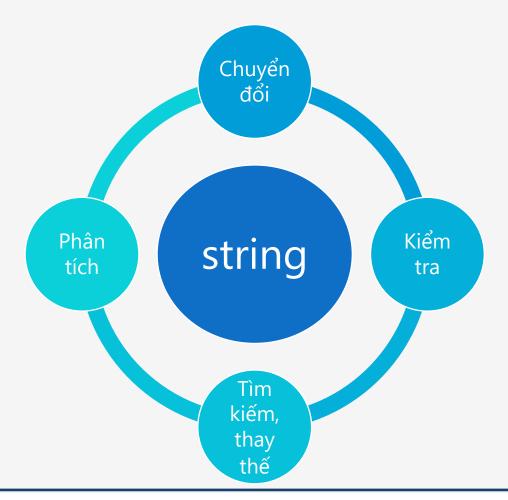
Hàm	Mô tả
len(st)	Trả về độ dài của xâu <i>st</i>
max(st)	Trả về kí tự lớn nhất (theo mã ASCII) trong xâu <i>st</i>
min(st)	Trả về kí tự nhỏ nhất (theo mã ASCII) trong xâu <i>st</i>

CÁC THAO TÁC VỚI XÂU KÍ TỰ



CÁC THAO TÁC VỚI XÂU KÍ TỰ

- Thực hiện dựa trên các phương thức của kiểu dữ liệu string
- Được chia làm 4 nhóm chính:
 - Chuyển đổi
 - Kiểm tra
 - Tìm kiếm, thay thế
 - Phân tích





THAO TÁC CHUYỂN ĐỔI

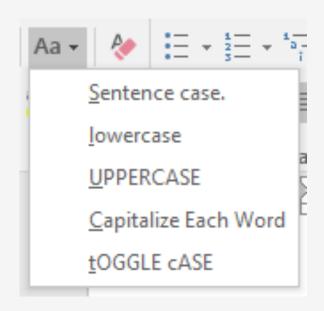
Phương thức	Mô tả
st.capitalize()	Viết hoa chữ cái đầu tiên của xâu
st.lower()	Trả về một bản sao của xâu <i>st</i> trong đó chữ hoa được chuyển sang chữ thường
st.swapcase()	Trả về một bản sao của xâu st trong đó kiểu chữ viết của tất cả các kí tự được đảo ngược, chữ hoa thành chữ thường và ngược lại
st.title()	Trả về một bản sao của xâu <i>st</i> trong đó tất cả kí tự đầu tiên của tất cả các từ là chữ hoa
st.upper()	Trả về một bản sao của xâu trong đó tất cả kí tự đều là chữ hoa



VÍ DŲ 7.3

Viết chương trình cho phép nhập vào từ bàn phím một chuỗi ký tự bất kỳ. Hãy mô phỏng lại một số chức năng xử lý chuỗi trong Microsoft Word như sau:

- Viết hoa (viết thường) tất cả các ký tự chữ cái của chuỗi.
- Viết hoa các chữ cái đầu câu
- Viết hoa các chữ cái đầu từ
- Viết thường các chữ cái đầu từ và viết hoa các chữ cái không phải đầu từ.





VÍ DŲ 7.3

```
s = "Cong Hoa xa hoi chu nghia"
st = s.lower()
print(st)
sh = s.upper()
print(sh)
ss = s.capitalize()
print(ss)
sw = s.title()
print(sw)
so = sw.swapcase()
print(so)
```

```
cong hoa xa hoi chu nghia
```



THAO TÁC KIỂM TRA

Phương thức	Mô tả
st.isalnum()	Trả về <i>True</i> nếu tất cả các kí tự trong xâu <i>st</i> là chữ cái hoặc chữ số, ngược lại trả về <i>False</i> .
st.isalpha()	Trả về <i>True</i> nếu xâu <i>st</i> có ít nhất 1 kí tự và tất cả các kí tự là chữ cái, ngược lại trả về <i>False</i> .
st.isdigit()	Trả về <i>True</i> nếu xâu <i>st</i> có ít nhất một kí tự và tất cả các kí tự trong xâu đều là chữ số, ngược lại, trả về <i>False</i> .
st.islower()	Trả về <i>True</i> nếu tất cả các kí tự trong xâu st là chữ thường, ngược lại trả về False.
st.isdecimal()	Trả về <i>Tru</i> e nếu một xâu chỉ chứa các kí tự số ở hệ thập phân, ngược lại trả về <i>Fal</i> se



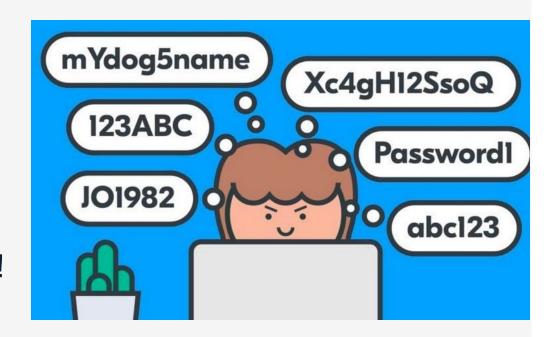
THAO TÁC KIỂM TRA

Phương thức	Mô tả
st.isnumeric()	Trả về <i>Tru</i> e nếu xâu <i>st</i> là dạng Unicode, chứa ít nhất một kí tự và tất cả các kí tự phải là kí tự số, ngược lại trả về <i>False</i> .
st.isspace()	Trả về <i>True</i> nếu xâu <i>st</i> chỉ chứa các kí tự khoảng trắng, ngược lại trả về <i>False</i> .
st.istitle()	Trả về <i>True</i> nếu xâu <i>st</i> có ít nhất một kí tự và tất cả kí tự phải ở dạng titlecase (kí tự đầu mỗi từ là chữ hoa, các kí tự khác là chữ thường), ngược lại trả về <i>False</i> .
st.isupper()	Trả về <i>True</i> nếu <i>st</i> có ít nhất một kí tự và tất cả kí tự trong xâu <i>st</i> là chữ hoa, ngược lại trả về <i>Fal</i> se.
st.startswith(substr [,start] [,stop])	Trả về <i>True</i> nếu xâu <i>st</i> hoặc xâu con từ <i>start</i> đến <i>stop</i> (nếu <i>start</i> và <i>stop</i> được cung cấp) bắt đầu với xâu con <i>substr</i> , ngược lại, trả về <i>False</i>
st.endswith(substr [,start] [,stop])	Trả về <i>True</i> nếu xâu <i>st</i> hoặc xâu con từ <i>start</i> đến <i>stop</i> (nếu <i>start</i> và <i>stop</i> được cung cấp) kết thúc với xâu con <i>substr</i> , ngược lại, trả về <i>False</i>





- Mật khẩu mạnh là mật khẩu thỏa mãn:
 - Có tối thiểu 8 kí tự
 - Có ít nhất một chữ cái thường
 - Có ít nhất một chữ cái hoa
 - Có ít nhất một chữ số
 - Có ít nhất một kí tự đặc biệt: @, #, \$, %, !
- Nhập một mật khẩu từ bàn phím. Kiểm tra xem đây có phải là mật khẩu mạnh hay không?







```
s = input("Nhap mat khau: ")
kt = ct = ch = cs = db = 0
for i in s:
   kt += 1
   if (i>='a' and i<='z'):</pre>
      ct += 1
   if (i>='A' and i<='Z'):
      ch += 1
   if (i>='0' and i<='9'):
      cs += 1
   if (i=='@' or i=='#' or i=='$' or i=='%'
or i=='!'):
      db += 1
print(kt, ct, ch, cs, db)
if (kt>=8 and ct>=1 and ch>=1 and cs>=1 and
db >= 1):
   print("Day la mat khau manh")
else:
   print("Day la mat khau yeu")
```

```
s = input("Nhap mat khau: ")
ct = ch = cs = db = 0
sdb = "@#$%!"
for i in s:
if (i.islower()):
      ct += 1
   if (i.isupper()):
      ch += 1
   if (i.isdigit()):
      cs += 1
   if (i in sdb):
      db += 1
print(len(s), ct, ch, cs, db)
if (len(s))=8 and ct>=1 and ch>=1 and cs>=1
and db >= 1):
        print("Day la mat khau manh")
else:
        print("Day la mat khau yeu")
```



TÌM KIẾM VÀ THAY THẾ

Phương thức	Mô tả
st.count(substr[, start][,	Đếm số lần xuất hiện của xâu <i>substr</i> trong xâu <i>st</i> hoặc xâu con của <i>st</i> nếu các tham
stop])	số start và stop được cung cấp. Mặc định start là đầu xâu và stop là cuối xâu.
st.find(substr[,start] [,stop])	Trả về vị trí xuất hiện đầu tiên của xâu <i>substr</i> trong xâu <i>st</i> hoặc xâu con từ vị trí <i>start</i> đến vị trí <i>stop</i> của xâu st. Nếu không tìm thấy, trả về -1.
st.replace(old, new [, max])	Trả về một bản sao của <i>st</i> trong đó thay thế tất cả các xâu <i>old</i> bằng xâu <i>new</i> với số lần tối đa là <i>max</i> . Tham số <i>max</i> có thể có hoặc không, nếu không mặc định là thay thế tất cả các xâu <i>old</i> tìm thấy.
st.rfind(substr [,start] [,stop])	Tương tự <i>find()</i> nhưng tìm kiếm từ phải sang trái, do đó trả về vị trí xuất hiện cuối cùng của xâu <i>substr</i> trong xâu <i>st</i> .
st.rindex(substr[,start] [,stop])	Tương tự <i>index()</i> ,nhưng tìm kiếm từ phải sang trái, do đó trả về vị trí xuất hiện cuối cùng của substr trong st nếu tìm thấy.

VÍ DŲ 7.5



- Họ tên của người Việt bao gồm ba thành phần: họ, tên đệm và tên. Hãy viết chương trình cho phép thực hiện các công việc sau:
 - Nhập vào họ tên đầy đủ của một người
 - Hiển thị họ của người đó ra màn hình
 - Hiển thị tên của người đó ra màn hình
 - Hiển thị họ đệm của người đó ra màn hình







```
s = input("Nhap ho ten: ")
i = s.find(" ")
ho = s[:i]
print(f"Ho: {ho}")
j = s.rfind(" ")
ten = s[(j+1):]
print(f"Ten: {ten}")
dem = s[(i+1):j]
print(f"Dem: {dem}")
```

```
Nhap ho ten: dang thanh trung
```

Ho: dang

Ten: trung

Dem: thanh

Nhap ho ten: dang thanh trung

Ho:

Ten:

Dem: dang thanh trung



PHÂN TÍCH

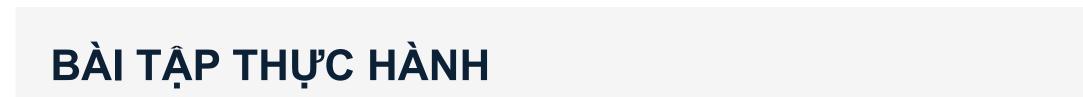
Phương thức	Mô tả
st.join(iter)	Trả về một xâu bằng cách nối các xâu là các phần tử trong dãy <i>iter</i> thành một xâu trong đó sử dụng <i>st</i> làm dấu phân tách giữa xâu con.
st.split(sep[,num])	Trả về danh sách các xâu con của st được chia theo sep. Nếu num được cung cấp, st được chia thành num + 1 xâu con theo sep.
st.splitlines([num])	Trả về một danh sách gồm tất cả các dòng trong xâu. Nếu <i>num</i> được cung cấp và là một số khác 0 thì dấu xuống dòng '\n' sẽ được thêm vào cuối mỗi xâu con.
st.join(iter)	Trả về một xâu bằng cách nối các xâu biểu diễn các phần tử trong dãy <i>iter</i> thành một xâu trong đó sử dụng <i>st</i> làm dấu phân tách giữa xâu con.
st.strip()	Trả về một xâu bằng cách loại bỏ các khoảng trắng ở bên trái và bên phải xâu
st.rstrip()	Trả về một xâu bằng cách loại bỏ các khoảng trắng ở bên trái xâu
st.lstrip()	Trả về một xâu bằng cách loại bỏ các khoảng trắng ở bên phải xâu





```
s = input("Nhap ho ten: ")
l = s.split()
print(1)
print(f"Ho: {1[0]}")
print(f"Ten: {1[2]}")
print(f"Dem: {1[1]}")
s1 = "-".join(1)
print(f"Ho ten: {s1}")
```

```
Nhap ho ten: dang thanh trung
['dang', 'thanh', 'trung']
Ho: dang
Ten: trung
Dem: thanh
Ho ten: dang-thanh-trung
```





Bài 7.3: Một số điện thoại gọi là hợp lệ nếu có 12 kí tự có dạng "+abcdefghijk". Trong đó a,b,c,d,e,f,g,h,I,j,k là các chữ số. Ví dụ: +84912675369, +84912345678

- Hãy viết chương trình thực hiện các công việc sau:
 - Nhập một chuỗi kí tự từ bàn phím
 - Kiểm tra xem một chuỗi đó có phải là một số điện thoại hợp lệ hay không?





Bài 7.4.

Hãy viết chương trình cho phép đổi định dạng ngày tháng "yyyy-mm-dd" thành định dạng "dd/mm/yyyy" và ngược lại. Ví dụ:

Nhap vao ngay thang: 1982-09-20

Dinh dang moi cua ngay thang: 20/09/1982





Bài 7.5. Microsoft Excel có một chức năng cho phép chèn dấu phân tách hàng nghìn cho một số bất kỳ. Hãy viết chương trình mô phỏng lại chức năng này như sau:

a. Cho phép nhập vào một chuỗi các chữ số

b. Thêm dấu phân cách hàng nghìn (dấu phẩy) vào số đó

Ví dụ:

```
N = "13324325684"
```

$$M = "13,324,325,684"$$

$$N = "842342235.5754"$$

$$M = "842,342,235.5754"$$



CÂU HỎI

