

TP n° 11

Chaînes de caractères

Rappels sur les chaînes de caractères

Notion de chaîne de caractères et type string (str)

Une chaîne de caractère est une suite finie de caractères entre apostrophes ('). Par exemple

```
'Informatik macht Spass'
```

est une chaîne de caractères. Nous disposons d'une chaîne de caractères vide : ''.

Le type correspondant aux chaînes de caractères est nommé string et noté str en Python. Afin d'illustrer quelques manipulations de chaînes de caractères, nous en stockons une dans une variable notée c, puis nous vérifions son type.

```
>>> c='Informatik macht Spass'  
>>> type(c)  
<class 'str'>
```

Longueur d'une chaîne de caractères

Nous pouvons obtenir la longueur d'une chaîne de caractères grâce à la fonction len(.).

```
>>> len(c)  
22
```

Indexation des caractères d'une chaîne de caractères

Les caractères d'une chaîne de caractères sont indicés de 0 à len(chaîne_de_caractères)-1. Nous accédons au caractère d'indice i d'une chaîne de caractère via

```
chaîne_de_caractères[i]
```

où i est un entier compris entre 0 et len(chaîne_de_caractères)-1 (au sens large). Si i est supérieur ou égal à len(chaîne_de_caractères), alors une erreur se produit.

```
>>> c[0]  
'I'  
>>> c[21]  
's'  
>>> c[19]  
'a'  
>>> c[22]  
Traceback (most recent call last):  
  File "<pyshell#17>", line 1, in <module>  
    c[22]  
IndexError: string index out of range
```

Nous pouvons aussi avoir accès à un caractère d'une chaîne de caractères « en comptant à partir de la fin » grâce à des indices négatifs, ceux-ci évoluant entre et -len(chaîne_de_caractères) et -1 (au sens large). Par exemple

```
chaîne_de_caractères[-1]
```

est le dernier caractère de la chaîne de caractères,

```
chaîne_de_caractères[-2]
```

est l'avant dernier caractère de la chaîne de caractères...

```
>>> c[-1]  
's'  
>>> c[-3]  
'a'  
>>> c[-22]  
'I'
```

```
>>> c[-23]
Traceback (most recent call last):
  File "<pyshell#25>", line 1, in <module>
    c[-23]
IndexError: string index out of range
```

Le slicing

Le slicing permet de prélever une suite de caractères (contigus ici) dans une chaîne de caractères donnée. Si i et j sont des indices compris entre 0 et $\text{len}(\text{chaîne_de_caractères})$ (au sens large) tels que $i < j$ alors

```
chaîne_de_caractères[i:j]
```

est la chaîne de caractères obtenue en prélevant les caractères d'indices compris entre i et $j-1$ (au sens large).

```
>>> c[1:4]
'nfo'
>>> c[2:3]
'f'
>>> c[19:22]
'ass'
>>> c[0:22]
'Informatik macht Spass'
```

Notons que des indices « trop grands », un ordre inversé entre les deux indices de prélèvement ne provoquent pas d'erreur. Ces modes d'utilisation du slicing rendent cependant les programmes difficiles à lire.

```
>>> c[0:100]
'Informatik macht Spass'
>>> c[4:2]
''
>>> c[25:100]
''
```

Mise en garde

Ces quelques rappels sur les chaînes de caractères ne sauraient se substituer au cours de M. Mollier (cf. partie 4 du chapitre 8), qui renferme bien d'autres notions, toutes aussi importantes (e.g. la concaténation + et les méthodes `append()`, `find()`, `count()`...).

Exercices

Exercice 1 (Comparaison des premier et dernier caractères d'une chaîne de caractères)

Écrire une fonction `first_last`

- d'argument c une chaîne de caractères;
- qui retourne `True` si la chaîne est vide ou si son premier caractère coïncide avec son dernier caractère, et `False` sinon.

Exercice 2 (Palindrome)

Écrire une fonction `palindrome`

- d'argument c une chaîne de caractères;
- qui retourne `True` si la chaîne est un palindrome¹, et `False` sinon.

Exercice 3 (Sous-chaîne de caractères présente dans une chaîne de caractères)

Écrire une fonction `include`

- d'arguments c_1 et c_2 deux chaînes de caractères;
- qui retourne `True` si la chaîne c_1 est une sous-chaîne de la chaîne c_2 , et `False` sinon.

Par exemple, si on exécute la fonction `include` avec comme arguments

```
c1='iseu' et c2='diviseur de tension'
```

elle renvoie `True`, et si on l'exécute avec comme arguments

```
c1='0hm' et c2='Loi des noeuds'
```

elle renvoie `False`. Pour réaliser la fonction `include`, nous pourrions utiliser le slicing, mais aucune méthode (en particulier ni `find()`, ni `count()`).

1. Un palindrome est une chaîne de caractères que l'on peut lire indifféremment de la gauche vers la droite ou de la droite vers la gauche. Par exemple 'LAVAL', 'NATAN', 'radar', 'tôt', 'mon nom' et 'éricnotrevaletallatelavertonciré' sont des palindromes.