

COURS : LISTES

Les programmes que nous avons écrits jusqu'à présent ne travaillaient qu'avec un nombre limité de valeurs. Nous avons par exemple écrit une fonction calculant le minimum de deux nombres flottants. Si notre algorithme doit travailler non pas avec deux nombres flottants, mais avec une dizaine de nombres, des milliers, voire des milliards, nous devons stocker cette collection de valeurs dans une *structure de données*. Différentes méthodes existent pour stocker en mémoire ces données. Selon l'algorithme que l'on souhaite utiliser pour traiter ces données, certaines sont plus performantes que d'autres. Dans ce chapitre, nous allons introduire la structure de donnée concrète la plus importante en informatique : *le tableau dynamique*. C'est avec la *liste chaînée*, la structure de donnée concrète la plus utilisée pour stocker une collection ordonnée d'objets. La première est fondamentale en programmation impérative alors que la seconde est fondamentale en programmation fonctionnelle. Python propose un tableau dynamique qu'il a eu le très mauvais goût d'appeler liste et qu'il ne faudra pas confondre avec la liste chaînée OCaml qui est aussi appelée liste, avec bon goût cette fois.

1 Liste

Une liste est une succession ordonnée de n valeurs. Si `a` est une liste de longueur n , ses valeurs sont indexées de 0 à $n - 1$ et il est possible d'accéder à la valeur d'indice k avec la notation `a[k]`. En mémoire, les valeurs sont stockées les unes à la suite des autres. On en déduit donc que si p est l'adresse mémoire de la valeur d'indice 0, la valeur d'indice k est située à l'adresse $p + k\delta$ où δ est le nombre d'octets utilisé pour stocker une valeur. Une liste a donc l'avantage d'avoir un accès direct à son k -ième élément. La longueur d'une liste s'obtient avec la fonction `len`.

```
1 >>> note = [9, 10, 14]
2 >>> note[0]
3 9
4 >>> (note[0] + note[1] + note[2]) / len(note)
5 11.0
```

Les éléments d'une sont accessibles en lecture et en écriture. Il est donc possible de changer leurs éléments. Si on veut changer la dernière note :

```
1 >>> note = [9, 10, 14]
2 >>> note[2] = 16
3 >>> note
4 [9, 10, 16]
```

Si on dépasse les bornes d'une liste, Python lance l'exception « list index out of range ». Par exemple `note[3]` va renvoyer une exception.

Si on souhaite ajouter un point à toutes les notes, il suffit d'effectuer une boucle sur l'indice de la liste.

```
1 >>> note = [9, 10, 14]
2 >>> for k in range(len(note)):
3     note[k] = note[k] + 1
4 >>> note
5 [10, 11, 15]
```

Nous avons vu comment créer une liste contenant quelques éléments. Mais si nous souhaitons créer une liste contenant n éléments, on utilise la notation suivante qui permet de créer ce que l'on appelle une liste en compréhension.

```
1 >>> n = 5
2 >>> a = [k**2 for k in range(n)]
3 >>> a
4 [0, 1, 4, 9, 16]
```

On peut aussi restreindre les éléments que l'on utilise en ajoutant une condition. On dit qu'on filtre la liste.

```
1 >>> a = [k**2 for k in range(n) if k % 3 != 1]
2 >>> a
3 [0, 4, 9]
```

Les listes sont dynamiques, c'est-à-dire qu'il est possible d'augmenter ou de diminuer leur taille une fois créées. Pour ajouter un élément en fin de liste, on utilise la méthode `append` qui va agir par effet de bord et ajouter 18 à la fin de la liste.

```
1 >>> note = [9, 10, 14]
2 >>> note.append(18)
3 >>> note
4 [9, 10, 14, 18]
```

Il est aussi possible d'enlever le dernier élément d'une liste. On utilise pour cela la méthode `pop` qui supprime le dernier élément de la liste et la renvoie.

```
1 >>> chapitre_du_cours_a_apprendre = [9, 8, 7, 6, 5]
2 >>> appris = chapitre_du_cours_a_apprendre.pop()
```

Les listes Python peuvent contenir des valeurs de types différents. Ils peuvent même contenir d'autres listes.

```
1 >>> a = [9, 3.14159, "Hello", True, []]
```

Cependant, en pratique, contrairement aux `tuple`, on n'utilisera que des listes contenant des valeurs ayant le même type.

2 Recherche dans une liste

Il est courant de se demander si une valeur est déjà présente dans une liste. Si l'on souhaite écrire un algorithme s'arrêtant dès que la valeur est trouvée, il y a deux manières courantes de programmer cet algorithme.

```
1 >>> def est_present_v1(x, lst):
2     for k in range(len(lst)):
3         if lst[k] == x:
4             return True
5     return False
```

Si on souhaite avoir un programme avec un seul point de sortie, on utilise plutôt la version suivante :

```
1 >>> def est_presentV2(x, lst):
2     k = 0
3     while k < len(lst) and lst[k] != x:
4         k = k + 1
5     return k != len(lst)
```

Cette version utilise le côté paresseux de l'opérateur `and`. En effet, cet opérateur possède une subtilité. Il évalue d'abord le terme à sa gauche et si il obtient `False`, il n'évalue pas le terme à sa droite, car on sait que le résultat de l'opération est `False`. Ici, si jamais la valeur cherchée ne se trouve pas dans la liste de longueur n , la variable `k` va prendre la valeur n et l'évaluation paresseuse de la condition va faire que la condition `k < len(lst)` (qui ne sera plus vérifiée) empêchera d'effectuer le test `lst[k] != x`, opération qui n'a aucun sens car l'index `k` est sorti des bornes de la liste. Au passage, l'opérateur `or` est aussi paresseux en Python : si son membre de gauche s'évalue en `True`, le membre de droite n'est pas évalué.

Les chaînes de caractères fonctionnant de manière similaire aux listes, voici quelques exercices de recherche sur ces dernières.

3 Recherche dichotomique dans une liste triée

Lorsqu'il existe une relation d'ordre totale sur les valeurs avec lesquelles on travaille et que la liste est *triée*, il est possible d'effectuer une recherche beaucoup plus rapide. Notons tout d'abord que si nous possédons un tableau, il est facile de le trier avec la méthode `sort`.

```
1 >>> liste = [8, 3, 7]
2 >>> liste.sort()
3 >>> liste
4 [3, 7, 8]
```

L'algorithme de recherche sur une liste triée que nous allons présenter est appelé algorithme de recherche dichotomique.

```
1 >>> def est_present_dichotomie(x, tab):
2     g = 0
```

```
3     d = len(tab) - 1
4     while g <= d:
5         m = (g + d) // 2
6         if x < tab[m]:
7             d = m - 1
8         elif x > tab[m]:
9             g = m + 1
10        else:
11            return True
12    return False
```

Remarquons au passage qu'il existe de nombreuses versions similaires de l'algorithme de recherche dichotomique. En voici une autre :

```
1 >>> def est_present_dichotomie_v2(x, tab):
2     g = 0
3     d = len(tab) - 1
4     while g < d:
5         m = (g + d) // 2
6         if x > tab[m]:
7             g = m + 1
8         else:
9             d = m
10    return tab[g] == x
```

4 Objets, types mutables et immutables

Les quelques expériences ci-dessous vont mettre à mal l'idée que l'on se fait des variables.

```
1 >>> a = [3, 9, 16]
2 >>> b = a
3 >>> a.append(25)
4 >>> a
5 [3, 9, 16, 25]
6 >>> b
7 [3, 9, 16, 25]
```

Ainsi, l'idée selon laquelle une variable est une boîte dans laquelle on place des valeurs est mise à mal. On voit en effet que les listes contenues dans les variables `a` et `b` ne sont pas indépendantes. Nous allons donc revenir sur certaines notions présentées dans les chapitres précédents qui ne correspondent pas à la réalité.

En Python, toutes les données sont représentées par des objets : les entiers, les nombres flottants, les booléens, les chaînes de caractères, les tuples, les listes et même les fonctions. Un objet possède un identifiant, un type et une valeur. L'identifiant d'un objet est l'adresse mémoire à laquelle est stocké l'objet. Sur un ordinateur 64 bits, c'est donc un entier compris entre 0 et $2^{64} - 1$. L'identifiant d'un objet ne change pas au cours de sa vie. De plus, deux objets distincts ont des identifiants différents. Un objet possède aussi un type qui ne change

pas au cours de la vie de l'objet. Enfin, un objet possède une valeur. Cette valeur peut ou non changer selon le type de l'objet. Les types élémentaires `bool`, `int`, `float` et `string` sont *immutable*. C'est aussi le cas des `tuple`. Pour les objets de types immutables, leur valeur ne peut pas changer. Le seul moyen d'avoir un une nouvelle valeur est de créer un nouvel objet. Les listes sont quant à elles de type *mutable*. Pour les objets de types `list`, leur valeur peut changer au cours de leur vie.

Contrairement à ce que nous avons annoncé, une variable n'est pas une boîte dans lequel on met une valeur. C'est plutôt une boîte dans laquelle on met l'identifiant de l'objet auquel il est associé. Par exemple, l'instruction

```
1 >>> a = 1
```

crée un objet de type `int` dont la valeur est 1 et associe le nom `a` à cet objet. En pratique, on peut voir une variable non pas comme une boîte contenant une valeur, mais comme une boîte contenant l'identifiant d'un objet. Lorsque l'on écrit

```
1 >>> a = a + 1
```

Python crée un nouvel objet de type `int` dont la valeur est 2 et associe le nom `a` à ce nouvel objet. La fonction `id`, donnant l'identifiant d'un objet nous permet d'observer cela.

```
1 >>> a = 1
2     print(id(a))
3     a = a + 1
4     print(id(a))
5 4326372464
6 4326372496
```

On voit qu'après avoir écrit `a = 1`, l'adresse mémoire de l'objet associé à `a` contenant la valeur 1 est différent de la seconde adresse mémoire qui est celle d'un objet qui contient la valeur 2. Lorsque l'on écrit

```
1 >>> a = 1
2     b = a
3     print(id(a))
4     print(id(b))
5 4326372464
6 4326372464
```

on voit que les variables `a` et `b` sont associées au même objet contenant la valeur 1. Mais puisque `int` est un type *immutable*, il est impossible de changer la valeur de cet objet. Il est donc impossible de changer la valeur associée à `b` en manipulant la variable `a`. Tout se passe donc comme si `a` et `b` étaient deux boîtes contenant directement la valeur 1.

Les choses changent si on travaille avec des types immutables. Si on écrit

```
1 >>> a = []
2     b = a
```

```
3     print(id(a))
4     print(id(b))
5 4547981064
6 4547981064
```

alors `a` et `b` sont associées à un même objet de type `list`, dont la valeur est une liste vide. Si on ajoute 1 à cette liste en utilisant la variable `a`, cet ajout se verra lorsque l'on travaille avec la variable `b`.

```
1 >>> a = []
2 >>> b = a
3 >>> a.append(1)
4 >>> b
5 [1]
```

Puisque `a` et `b` sont associées au même objet, on dit que ces variables présentent un *aliasing*.

Tout comme une variable, une liste ne contient pas des valeurs, mais des identifiants d'objets. Il est donc possible d'avoir un aliasing entre différents éléments d'un tableau. Par exemple :

```
1 >>> a = []
2 >>> b = [a, a]
3 >>> b
4 [[], []]
5 >>> a.append(42)
6 >>> b
7 [[42], [42]]
```

Lorsque l'on crée une liste en compréhension, ce sont des objets distincts qui sont associés au éléments d'un tableau.

```
1 >>> n = 3
2 >>> a = [[] for k in range(3)]
3 >>> a
4 [[], [], []]
5 >>> a[0].append(42)
6 >>> a
7 [[42], [], []]
```

En ce qui concerne les `tuple`, leur caractère immutable signifie que les identifiants des objets le formant ne peuvent pas être changés. Si ces objets sont mutables, il est possible de changer leur valeur.

```
1 >>> t = ([], [])
2 >>> t[0].append(3)
3 >>> t
4 ([3], [])
```

Enfin ce n'est pas une valeur que l'on passe à une fonction, mais l'identifiant d'un objet. Si la fonction modifie cet objet, elle aura un effet sur l'objet passé à la fonction. On dit alors que la fonction agit par effet de bord. Par exemple :

```

1  >>> def f(lst):
2      lst.append(42)
3  >>> a = []
4  >>> f(lst)
5  >>> a
6  [42]

```

Par contre, si on associe la variable locale à un nouvel objet, cela n'aura aucune influence sur l'objet initial.

```

1  >>> def f(lst):
2      lst = [42]
3  >>> a = []
4  >>> f(lst)
5  >>> a
6  []

```

On dit que le passage de paramètre se fait *par objet* en Python. Notons que selon les communautés, le nom utilisé pour ce type de passage diffère. Mais c'est inutile de lancer une nouvelle guerre informatique : le nom donné à ce type de passage diffère d'une communauté à l'autre. L'essentiel est de savoir que les personnes qui n'utilisent pas le terme de « passage par objet » ont tort. La dernière phrase est bien sûr à prendre au second degré.

En résumé, on retiendra les points essentiels suivants :

- En Python, les données sont représentées par des *objets*.
- Certains objets sont *immuables*. Ce sont les objets de type `bool`, `int`, `float`, `str`, `tuple`. D'autres objets sont de type *mutable*. C'est le cas par exemple du type `list`.
- Si on travaille avec des objets immuables, tout se passe comme si une variable était une boîte contenant la valeur d'un objet. Cependant, avec des objets mutables, il est essentiel de réaliser qu'une variable contient non pas une valeur, mais l'identifiant d'un objet. Deux variables peuvent contenir le même identifiant et donc offrir un accès au même objet. On dit dans ce cas qu'il y a *aliasing*.
- En Python, le passage des paramètres se fait *par objet*. En pratique, avec des types immuables, tout se fait comme si on copiait la valeur. Mais avec les types mutables, il est essentiel de réaliser que c'est l'identifiant de l'objet et non sa valeur qui est passée à la fonction. Quand on passe un objet mutable à une fonction, cette dernière peut donc potentiellement le modifier. Si c'est le cas, on dit que la fonction agit par effet de bord.

5 Concaténation, Slicing

Comme pour les chaînes de caractères, il est possible de concaténer deux listes. On peut même multiplier des listes par un entier.

```

1  >>> [2, 3, 4] + [5, 6, 7]
2  [2, 3, 4, 5, 6, 7]
3  >>> [1, 2, 3] * 3
4  [1, 2, 3, 1, 2, 3, 1, 2, 3]

```

Faites cependant très attention à ces opérations qui sont très dangereuses dès lors que votre liste contient des objets mutables. Étant donné qu'une liste, tout comme une variable, contient en fait les identifiants des objets, on peut avoir quelques surprises dues à l'aliasing. C'est par exemple une très mauvaise idée de construire une liste de listes vides de cette manière.

```

1  >>> n = 2
2  >>> liste = [[]] * 2
3  >>> liste[0].append(0)
4  >>> liste
5  [[0], [0]]

```

On préférera utiliser la compréhension qui provoque moins de soucis :

```

1  >>> n = 2
2  >>> liste = [[] for k in range(n)]
3  >>> liste.append(0)
4  >>> liste
5  [[0], []]

```

Il est possible de construire des sous-listes avec une opération que l'on appelle le slicing.

```

1  >>> liste = ["Zero", "Un", "Deux"]
2  >>> a = liste[0:2]
3  >>> a
4  ["Zero", "Un"]

```

Plus généralement, si `liste` est de longueur n et $(a, b) \in \llbracket 0, n \rrbracket^2$ sont tels que $a \leq b$, alors `liste[a:b]` est la sous-liste composée des objets `liste[a]`, ..., `liste[b-1]`. Encore une fois, méfiez-vous de l'aliasing si vous travaillez avec des types mutables :

```

1  >>> liste = [[], [0], [0, 1]]
2  >>> sous_liste = liste[0:2]
3  >>> sous_liste[0].append(3)
4  >>> sous_liste
5  [[3], [0]]
6  >>> liste
7  [[3], [0], [0, 1]]

```

Enfin, Python possède une particularité qui permet d'accéder facilement aux listes par la fin. En effet, l'index -1 peut être utilisé pour accéder au dernier élément, l'index -2 à l'avant-dernier, etc. Par exemple :

```

1  >>> liste = ["Zero", "Un", "Deux"]
2  >>> liste[-1]
3  "Deux"
4  >> liste[1:-1]
5  ["Zero", "Un"]

```

Les mêmes opérations, excepté les modifications, sont disponibles pour les chaînes de caractères.

6 Types itérables

Les listes sont des objets itérables, c'est-à-dire qu'il est possible d'écrire des boucles `for` où l'indice prend les valeurs successives de la liste.

```
1  >>> liste = [2, 3, 5, 7, 11]
2  >>> somme = 0
3      for i in liste:
4          somme = somme + i
5  >>> somme
```

Il existe de nombreux types Python qui sont itérables. C'est le cas par exemple des chaînes de caractères. Un peu comme un croupier qui distribue les cartes, les types itérables vous

donnent les éléments un à un.

Si votre liste contient des tuples, il est possible d'utiliser plusieurs variables dans la boucle. Par exemple :

```
1  >>> liste = [("Lyon", 515695), ("Marseille", 869815), ("Paris", 2140526)]
2  >>> for ville, population in liste:
3      print(ville, "a", population, "habitants.")
```

On peut aussi utiliser la fonction `enumerate` sur une liste pour obtenir l'effet suivant :

```
1  >>> notes = [6, 12, 13]
2  >>> for k, note in enumerate(notes):
3      print("Au devoir", k + 1, "ma note est", note + ".")
```