Classes et POO

Définition d'une classe

- Une classe simple est introduite par le mot-clé *class*, suivi du nom de la classe :

```
class NomClasse:
  ... corps de la classe : définition de méthodes, affectation de variables.
```

- Attention, par défaut les classes Python sont plutôt permissives (tout est public, on peut rajouter des champs à un objet, etc.) :

```
class Cercle:
  pass

un_cercle = Cercle()   # Crée un objet Cercle
un_cercle.rayon = 5    # On peut lui rajouter un attribut à la volée !
del un_cercle.rayon    # Ou en supprimer...
```

- Généralement, on ajoute la méthode spéciale *__init__* pour initialiser une instance lors de sa création (celle-ci joue donc le rôle des « constructeurs » de C++ ou Java) :

```
class Cercle:
  def __init__(self, rayon):   # Toutes les méthodes doivent avoir self comme premier paramètre
    self.rayon = rayon         # self.rayon est une variable d'instance, rayon est le paramètre

un_cercle = Cercle(5)   # Appel implicite de Cercle.__init__(un_cercle, 5)
```

Définition

Quelques remarques importantes à savoir lorsque l'on a pratiqué d'autres langages de POO :

- Les méthodes doivent avoir un premier paramètre *self* qui désignera l'instance au moment de l'appel. Ce paramètre est passé implicitement à l'appel de la méthode.
- Les variables d'instance doivent obligatoirement être préfixées par *self* afin de les distinguer des variables locales et des paramètres. Python étant un langage dynamique, les variables d'instances sont créées lors de leur première affectation.
- Par défaut, les méthodes et les variables sont publiques et les classes sont « ouvertes » (on peut leur rajouter/ôter des variables et des méthodes depuis l'extérieur).
- Mais le programmeur peut créer des méthodes et des variables privées, empêcher la modification d'une classe, etc.

Variables d'instance et variables de classe

- Dans le corps de la classe, les variables d'instances sont préfixées par *self* et sont créées lors de leur première affectation (généralement, dans *__init__*).
- Toute variable initialisée en dehors de toute méthode est considérée comme une variable de classe : elle existe même si aucune instance n'a été créée et sa valeur est partagée par toutes les instances de la classe.
- Pour accéder à une variable de classe, il faut préfixer son nom du nom de la classe.
- Attention à certains problèmes (voir l'exemple).

```
class Cercle:
  pi = 3.14159                # Variable de classe
  def __init__(self, rayon):  # Toutes les méthodes doivent avoir self comme premier paramètre
    self.rayon = rayon        # self.rayon est une variable d'instance, rayon est le paramètre

print(Cercle.pi)              # Affiche 3.14159
c = Cercle(2)                 # Crée un cercle de rayon 2
print(c.rayon)                # Affiche 2
print(c.pi)                   # Affiche 3.14159
c.pi = 12                     # Ouch : on vient de créer une variable d'instance pi pour l'objet c
print(c.pi)                   # Affiche 12...
print(Cercle.pi)              # Affiche 3.14159
```

Variables d'instance et variables de classe

Si l'on veut vraiment empêcher la création dynamique d'attributs d'instance, on peut utiliser la méthode spéciale `__setattr__` :

```
class Cercle:
  pi = 3.14159                # Variable de classe
  def __init__(self, rayon):  # Toutes les méthodes doivent avoir self comme premier paramètre
    self.rayon = rayon        # self.rayon est une variable d'instance, rayon est le paramètre

  def __setattr__(self, nom, val):  # Appelée à chaque affectation d'un attribut
    if nom not in ['rayon']:
       raise AttributeError(f"{nom} n'est pas défini")
    else:
       self.__dict__[nom] = val

c = Cercle(2)
c.pi = 12                     # AttributeError: pi n'est pas défini
c.pouet = 24                  # AttributeError: pouet n'est pas défini
c.rayon = 42                  # Ok
```

Mais ce n'est que rarement nécessaire...

Méthodes d'instance et de classe

- Alors que les méthodes d'instance s'appliquent à une instance particulière de la classe, les méthodes de classe s'appliquent à *toutes* les instances d'une classe.
- Python permet de créer à la fois des méthodes *de classe* et des méthodes *statiques*. Leur principale différence est la syntaxe de leur définition.
- Une méthode de classe ou une méthode statique peut être appelé sur le nom de la classe où sur celui d'une instance.
- Les décorateurs *@classmethod* et *@staticmethod* permettent, respectivement, de définir une méthode de classe et une méthode statique.
- Évidemment, une méthode statique ou de classe qui manipulerait une variable d'instance n'aurait aucun sens (alors qu'une méthode d'instance peut manipuler une variable de classe).

Exemple de méthode statique

Fichier cercle.py

```
class Cercle:
  # Variables de classe
  tous = []
  pi   = 3.14159

  def __init__(self, rayon):
     self.rayon = rayon
     Cercle.tous.append(self)   # On ajoute le cercle à la liste de tous les cercles
                                # Ou : self.__class__.tous.append(self)
  def surface(self):
     return Classe.pi * self.rayon**2

  @staticmethod
  def surface_totale():
     total = 0
     for c in Cercle.tous:
        total += c.surface()
     return total
```

Exemple d'utilisation :

```
import cercle
c = cercle.Cercle(4)
c2 = cercle.Cercle(3)
print(cercle.Cercle.surface_totale())        # 78.53975
```


Exemple de méthode de classe

Fichier cercle.py

```
class Cercle:
  ... idem précédemment ...

  @classmethod
  def surface_totale(cls):
     total = 0
     for c in cls.tous:
        total += c.surface()
     return total
```

- L'avantage d'utiliser une méthode de classe est que cette méthode connaît la classe via le paramètre qui lui a été passé (*cls*, ici). On peut donc écrire simplement *cls.tous* dans le corps de la méthode.
- Une méthode statique n'a pas ce paramètre et doit donc soit coder « en dur » le nom de la classe (*Cercle.tous*, dans notre exemple), soit passer par la variable spéciale *__class__*.
- Une méthode de classe peut également être redéfinie dans une sous-classe, pas une méthode statique... Mon conseil : utilisez plutôt des méthodes de classe.

Variables et méthodes privées

- Par défaut, tous les membres d'une classe sont *publics* : les variables d'instance notamment, peuvent être modifiées depuis l'extérieur, ce qui nuit au principe d'encapsulation des données.
- En Python, il n'existe pas de mot-clé *private* comme en Java, C++ ou C# : pour créer des variables et des méthodes privées, il suffit de préfixer leur nom par deux blancs soulignés.
- En réalité, l'attribut (ou la méthode) ne sont pas « privés » : leur nom est simplement modifié par Python pour compliquer leur accès depuis l'extérieur de la classe.
- Attention : un nom de membre privé ne doit pas se terminer par deux blancs soulignés (ce format est réservé aux noms spéciaux de Python, comme *__init__* ou *__doc__*).
- L'avantage de cette convention est qu'elle facilite l'identification des membres privés.

Exemple

Fichier point.py

```
import math

class Point:
  "Classe définissant un point du plan"

  def __init__(self, x, y):
    "Crée le point d'abscisse x et d'ordonnée y"
    self.__x, self.__y = x, y                    # __x et __y sont "privées"

  def getX(self): return self.__x
  def getY(self): return self.__y

  def __distance_origine(self):                  # __distance_origine est "privée"
     return math.hypot(self.__x, self.__y)

  def getDistance(self):
     "Renvoie la distance du point par rapport à l'origine (0,0)"
     return self.__distance_origine()
```

Exemple d'utilisation :

```
from point import Point
p = Point(3, 2)
p.__x                   # AttributeError: 'Point' object has no attribute '__x'
p.getX()                # 3
p.__distance_origine()  # AttributeError: 'Point' object has no attribute '__distance_origine'
p.getDistance()         # 3.605551275463989
```

Propriétés

- Au lieu d'utiliser des accesseurs Java-esque comme *getX()* ou *setX(valeur)* pour lire ou modifier des variables d'instances privées, il est préférable d'utiliser des *propriétés*.
- L'avantage des propriétés est qu'elles allègent le code utilisateur et qu'elles assurent le principe d'accès uniforme : l'utilisateur d'une classe n'a pas besoin de savoir si une information lui est fournie par une méthode ou par une variable.
- Ce mécanisme est également utilisé par d'autres langages, comme C# ou Ruby.
- Une propriété de lecture est introduite par le décorateur *@property*.
- Si l'on veut lui ajouter une propriété d'écriture, on ajoute à cette propriété un décorateur « setter » (voir exemple).

Exemple

```
import math

class Point:
  """Classe définissant un point du plan"""

  def __init__(self, x, y):
    """Crée le point d'abscisse x et d'ordonnée y"""
    self.__x, self.__y = x, y

  @property
  def x(self): return self.__x
  @x.setter
  def x(self, new_x): self.__x = new_x

  @property
  def y(self): return self.__y
  @y.setter
  def y(self, new_y): self.__y = new_y

  @property
  def distance(self): return math.hypot(self.__x, self.__y)
```

Exemple d'utilisation :

```
from point import Point
p = Point(3, 2)
p.x                     # 3 (utilisation de la propriété en lecture x)
p.y = 4                 # utilisation de la propriété en écriture y
p.distance              # 5.0 (utilisation de la propriété en lecture distance)
```

Héritage

- Une classe Python peut hériter d'une ou plusieurs classes. En réalité, les classes précédentes héritaient de *object*, la classe racine de la hiérarchie des classes Python.
- Contrairement à Java, Ruby ou C#, Python autorise l'héritage multiple : une classe peut hériter de plusieurs classes.
- Une classe hérite de tous les membres non privés de ses super-classes et elle peut redéfinir certaines méthodes.
- Un appel à *super().une_methode()* permet d'appeler la méthode *une_methode()* définie dans une super-classe (mais son utilisation est critiquée par certains, **notamment par moi**...).
- L'ordre de recherche des attributs part de la classe, puis remonte dans la première super-classe de la liste, puis remonte dans les super-classes de celle-ci. La recherche se poursuit avec la seconde super-classe, etc. On a donc une recherche de gauche à droite, en profondeur d'abord.
- La fonction *isinstance(obj, cls)* teste si un *obj* est une instance de *cls* ou de l'une de ses sous-classes.
- La fonction *issubclass(cls1, cls2)* teste si la classe *cls1* est une sous-classe de *cls2*.

Exemple

```
class PointNomme(Point):
  """Classe définissant un point étiqueté"""

  def __init__(self, x, y, nom):
     Point.__init__(self, x, y)            # Appel du constructeur de Point
     # Ou : super().__init__(x, y)
     self.__nom = nom

  @property
  def nom(self): return self.__nom
```

Exemple d'utilisation :

```
import point2

p = point2.PointNomme(3, 2, "Point 1")
p.nom                                  # 'Point 1'
p.x                                    # 3
p.distance                             # 3.605551275463989
p.y = 3
p.distance                             # 4.242640687119285
p.__class__                            # <class 'point2.PointNomme'>
```

Polymorphisme

Toutes les méthodes d'une classe Python sont virtuelles, ce qui signifie qu'elles peuvent être redéfinies dans les classes filles :

```
class ClasseBase:

  def __init__(self, x, y):
    self.x, self.y = x, y

  def deplace(self, delta_x, delta_y):
     self.efface()
     self.x += delta_x
     self.y += delta_y
     self.affiche()

  def efface(self):   # on ne sait pas encore le faire...
    pass              # Ou  ...  (en Python 3)

  def affiche(self):  # on ne sait pas encore le faire...
    pass

class Fille(ClasseBase):

   def __init__(self, x, y, nom):
      ClasseBase.__init__(self, x, y)      # Ou super().__init__(x, y)
      self.nom = nom

   def efface(self):
      print(f"{self.nom} s'efface...")

   def affiche(self):
      print(f"{self.nom} s'affiche...")
```


Polymorphisme

Exemple d'utilisation :

```
import polymorph

f = polymorph.Fille(10, 12, "fille")
f.x                                    # 10
f.y                                    # 12
f.deplace(10, 10)                      # deplace() est définie dans ClasseBase
                                       # Affiche 'fille s'efface...'
                                       # Affiche 'fille s'affiche...'
f.x                                    # 20
f.y                                    # 22
```

- L'appel *f.deplace(...)* a appelé la méthode *deplace()* de la super-classe.
- Comme la classe *Fille* redéfinit les méthodes *efface()* et *affiche()*, la méthode *deplace()* les appelle car ce sont elles qui sont « les plus proches » de l'objet *f*.
- On a donc un « double dispatch » : comme *Fille* ne définit pas *deplace()*, on remonte vers sa super-classe pour trouver *deplace()*, puis on redescend vers la classe de *f* pour trouver les méthodes *efface()* et *affiche()* les plus spécialisées.
- Techniquement, on n'a pas besoin de définir les méthodes *efface()* et *affiche()* dans *ClasseBase*. Mais, si on ne le fait pas, un appel à *deplace()* sur un objet de *ClasseBase* provoquera un *AttributeError*. Pour simuler une *classe abstraite*, on peut également implémenter ces deux méthodes pour qu'elles renvoient *NotImplemented*.

Représentation textuelle

- Les deux méthodes spéciales *__repr__* et *__str__* permettent de gérer la représentation textuelle d'un objet. Ces deux méthodes sont appelées automatiquement par les fonctions *repr()* et *str()*.
- Par défaut, toutes les deux produisent une chaîne décrivant l'objet : *<point3.Point object at 0x7f5a2f857910>*, par exemple.
- La méthode *__repr__* est censée produire une représentation textuelle de l'objet. Cette représentation n'est pas destinée à être lue par un humain, mais doit permettre de recréer l'objet via un appel à *eval()* (cf. exemple). Elle est appelée automatiquement par la fonction *repr(obj)* (et par idle ou l'interpréteur intéractif).
- La méthode *__str__* est censée produire une représentation lisible de l'objet. Elle est appelée automatiquement par les instructions comme *print* et par la fonction de conversion *str(obj)*.

Exemple

```
# Module point4.py

class Point:
  (...)
  def __repr__(self):
    return f"Point({self.x}, {self.y})"

  def __str__(self):
    return f"({self.x}, {self.y})"

class PointNomme(Point):
  (...)
  def __repr__(self):
    return f"PointNomme({self.x}, {self.y}, {self.nom})"

  def __str__(self):
    return f"({self.x}, {self.y}, {self.nom})"
```

Exemple d'utilisation :

```
from point4 import Point

p = Point(3, 2)
repr(p)                                 # Point(3, 2)
q = eval(repr(p))                       # exécute donc q = Point(3, 2)...
r = eval(p.__module__ + '.' + repr(p))  # Si on avait simplement fait 'import point4'
q == p                                  # True
print(p)                                # Affiche (3, 2)  (appel de str(p))
```

Comparaisons

- Par défaut, deux objets *Point* distincts seront toujours considérés comme différents (l'égalité par défaut compare les références...).
- En réalité, les opérateurs de comparaison (et les autres aussi, d'ailleurs) appellent des méthodes spéciales d'*object* que l'on peut redéfinir pour modifier leurs comportements.
- Dans le cas des opérateurs de comparaison, il s'agit des méthodes *__eq__*, *__ne__*, *__lt__*, *__le__*, *__gt__* et *__ge__*.
- Si l'on redéfinit *__eq__*, Python redéfinit automatiquement *__ne__* si on ne le fait pas (mais autant le faire explicitement).
- Pour les comparaisons, une bonne pratique consiste à redéfinir tous les opérateurs afin qu'il n'y ait pas d'incohérences. Afin de faciliter, cette tâche, on dispose de *functools.total_ordering* qui les génère tous à partir de *__eq__* et de l'un des opérateurs *__lt__*, *__le__*, *__gt__* et *__ge__*.
- L'instruction *obj1 == obj2* appelle en réalité *obj1.__eq__(obj2)* (idem pour les autres fonctions de comparaison).

Exemple

```
import math
from functools import total_ordering

@total_ordering
class Point:
  def __init__(self, x, y):
    self.__x, self.__y = x, y

  def __eq__(self, autre):
    if not isinstance(autre, Point):
      return NotImplemented          # Python essaiera de trouver autre.__eq__(self)
    else:
      return self.__x == autre.x and self.__y == autre.y

  def __lt__(self, autre):
    if not isinstance(autre, Point):
      return NotImplemented          # Python essaiera de trouver autre.__lt__(self)
    else:
      return self.distance < autre.distance

  @property
  def x(self): return self.__x
  @x.setter
  def x(self, new_x): self.__x = new_x

  @property
  def y(self): return self.__y
  @y.setter
  def y(self, new_y): self.__y = new_y

  @property
  def distance(self): return math.hypot(self.__x, self.__y)
```

Exemple (suite)

```
@total_ordering
class PointNomme(Point):
  """Classe définissant un point étiqueté"""

  def __init__(self, x, y, nom):
     Point.__init__(self, x, y)         # Ou super().__init__(x, y)
     self.__nom = nom

  def __eq__(self, autre):
    return Point.__eq__(self, autre) and self.__nom == autre.nom
    # ou return super().__eq__(autre) and self.__nom == autre.nom

  def __lt__(self, autre):
    return Point.__lt__(self, autre) and self.__nom < autre.nom
    # ou return super().__lt__(autre) or self.__nom < autre.nom

  @property
  def nom(self): return self.__nom
```

Tests :

```
p = point3.Point(3, 2)
q = point3.Point(3, 2)
r = point3.Point(4, 6)
p == q                          # True
p != r                          # True
p < r                           # True
p > r                           # False
p <= q                          # True
```


La méthode spéciale __hash__

- On rappelle que pour pouvoir servir de clé, un objet doit être immutable et disposer d'une méthode *__hash__* (qui est appelée automatiquement par la fonction *hash(obj)*).
- Les objets *Point* ne sont pas immutables (pour qu'ils le soient, il suffit de supprimer les deux propriétés setters qui permettent de modifier les attributs du point). Voir l'exemple suivant.
- Si l'on redéfinit la méthode spéciale *__eq__*, Python ne fournit plus de méthode *__hash__* par défaut : les objets de notre classe ne peuvent donc plus servir de clé de dictionnaire.
- La création d'une méthode de hachage est un problème épineux car il faut s'assurer qu'elle fournira une valeur différente pour chaque objet différent au sens de *__eq__*.
- Dans le cas d'un point, notamment, il faut que le hachage d'un *Point(3, 2)* soit différent du hachage d'un *Point(2, 3)* et que *tous* les *Point(3, 2)* renverront la même valeur de hachage.
- Pour écrire une nouvelle fonction de hachage, le plus simple (et le plus sûr) consiste à utiliser une fonction de hachage existante.

Exemple

```
# Module point5.py

import math
from functools import total_ordering

@total_ordering
class Point:
  """Classe définissant un point du plan"""
  (...)
  @property
  def x(self): return self.__x      # On a supprimé le setter

  @property
  def y(self): return self.__y      # On a supprimé le setter

  @property
  def distance(self): return math.hypot(self.__x, self.__y)

  def __hash__(self):
   return hash((self.__x, self.__y))  # le hachage du tuple (x,y) sera différent de celui de (y, x)

@total_ordering
class PointNomme(Point):
  """Classe définissant un point étiqueté"""
  (...)
  @property
  def nom(self): return self.__nom

  def __hash__(self):
    return hash((self.__x, self.__y, self.__nom))
```


Exemple

La fonction prédéfinie *hash(obj)* appelle la méthode *obj.__hash__()* :

```
from point5 import Point

p = Point(3, 2)
q = Point(3, 2)
r = Point(2, 3)

id(p)                       # 139751287446864
id(q)                       # 139751287446672 (donc différent alors que p == q)
hash(p)                     # 3713083796995235906
hash(q)                     # 3713083796995235906 (donc égal)
hash(r)                     # 3713082714463740756

d = {p: "toto", r: "titi"}  # {Point(3, 2): 'toto', Point(2, 3): 'titi'}
d[q] = "normal"
d                           # {Point(3, 2): 'normal', Point(2, 3): 'titi'}
p = Point(1, 1)
d                           # {Point(3, 2): 'normal...', Point(2, 3): 'titi'}
d[p] = "autre"
d                           # {Point(1, 2): 'autre', Point(3, 2): 'normal...', Point(2, 3): 'titi'}
```

Remarque : la fonction *id(obj)* renvoie l'adresse mémoire de l'objet concerné. C'est donc un moyen d'identifier de façon unique deux objets différents mais elle considérera aussi comme différents deux points qui ont pourtant les mêmes coordonnées et qui, du point de vue d'un hachage, devraient être considérés comme égaux.

Création d'une classe collection

- Le but, ici, est de créer une classe « collection » qui se comporte comme les classes collections prédéfinies (*list*, *dict*, *tuple*)...
- On veut notamment pouvoir accéder en lecture (et en écriture, si notre classe n'est pas immutable) à un élément en utilisant la notation des crochets et on veut pouvoir parcourir tous les éléments de notre collection à l'aide d'une boucle *for*.
- Les méthodes spéciales *__getitem__*, *__setitem__* et *__delitem__* permettent de bénéficier de la notation entre crochets.
- La méthode *__contains__* permettent d'utiliser les opérateurs *in* et *not in* pour tester la présence d'un élément dans une collection.
- Les méthodes spéciales *__add__*, *__mul__* et *__len__* permettent d'implémenter, respectivement, la concaténation, la multiplication et la longueur d'une collection.
- La méthode spéciale *__iter__* permet de parcourir les éléments d'une collection. Elle est appelée automatiquement par *for*.

Premier exemple : un type ensemble

```
class Ensemble:
  """ Implémentation d'un ensemble d'entiers à l'aide d'un hachage
      dont les clés sont les entiers et les valeurs des booléens """

  def __init__(self, *liste):
     self.__corps = {}
     for e in liste: self.__corps[e] = True

  def contient(self, elt):
    return self.__corps.get(elt, False)   # ou : return elt in self.__corps.keys()

  def ajoute(self, elt):
    self.__corps[elt] = True

  def supprime(self, elt):
    if self.__corps.get(elt): del self.__corps[elt]

  def __str__(self):
    s = ""
    for e in self.__corps.keys(): s += str(e) + ", "
    return s[:-2]

  def union(self, autre):
    if not isinstance(autre, Ensemble): return self
    res = Ensemble()
    for e in self.__corps.keys(): res.ajoute(e)
    for e in autre.__corps.keys(): res.ajoute(e)
    return res

  (...)
```

Premier exemple : un type ensemble

```
  def intersect(self, autre):
    if not isinstance(autre, Ensemble): return None
    res = Ensemble()
    for e in self.__corps.keys():
       if e in autre.__corps.keys():
         res.ajoute(e)
    return res

  def diff(self, autre):
    if not isinstance(autre, Ensemble): return self
    res = Ensemble()
    for e in self.__corps.keys():
       if e not in autre.__corps.keys():
         res.ajoute(e)
    return res

if __name__ == '__main__':
  ens = Ensemble(12, 2, 1, 3)
  print(ens)                          # 1, 2, 3, 12
  ens.ajoute(30)
  print(ens)                          # 1, 2, 3, 12, 30
  ens.supprime(100)
  ens.supprime(30)
  print(ens)                          # 1, 2, 3, 12
  assert ens.contient(3)
  assert not ens.contient(42)
  ens2 = Ensemble(3, 2, 100, 34)
  print(ens2)                         # 2, 3, 100, 34
  print(ens.union(ens2))              # 1, 2, 3, 100, 12, 34
  print(ens.intersect(ens2))          # 2, 3
  print(ens.diff(ens2))               # 1, 12
```


Premier exemple : un type ensemble

Ça marche, mais c'est assez « Java-esque »... Voici une nouvelle version, plus « pythonique » :

```
class Ensemble:

  def __init__(self, *liste):
     self.__corps = {}
     for e in liste: self.__corps[e] = True

  def __iadd__(self, elt):
     self.__corps[elt] = True
     return self

  def __str__(self):
    s = ""
    for e in sorted(self.__corps.keys()): s += str(e) + ", "
    return s[:-2]

  def __or__(self, autre):
    if not isinstance(autre, Ensemble): return self
    res = Ensemble()
    for e in self: res += e   # Utilisation de __contains__ et de __iadd__
    for e in autre: res += e  # idem
    return res

(...)
```

Premier exemple : un type ensemble

```
def __and__(self, autre):
  if not isinstance(autre, Ensemble): return None
  res = Ensemble()
  for e in self:            # Utilisation de __contains__
     if e in autre:         # Utilisation de __contains__
       res += e             # Utilisation de __iadd__
  return res

def __sub__(self, autre):
  if not isinstance(autre, Ensemble): return self
  res = Ensemble()
  for e in self:
     if e not in autre:
       res += e
  return res

def __iter__(self):
  for e in sorted(self.__corps.keys()): yield e

def __getitem__(self, e):
  return self.__corps.get(e, False)

def __contains__(self, e):
  return self[e]

def __delitem__(self, e):
  self.__corps.pop(e, None)

def __len__(self):
  return len(self.__corps)
```


Premier exemple : un type ensemble

Et l'on peut maintenant écrire :

```
if __name__ == '__main__':
  ens = Ensemble(12, 2, 1, 3)
  print(ens)                                    # 1, 2, 3, 12
  ens += 30
  print(ens)                                    # 1, 2, 3, 12, 30
  del ens[100]
  del ens[30]
  print(ens)                                    # 1, 2, 3, 12
  assert ens[3]
  assert (3 in ens)
  assert not ens[42]
  assert (42 not in ens)
  ens2 = Ensemble(3, 2, 100, 34)
  print(ens2)                                   # 2, 3, 34, 100
  print(ens | ens2)                             # 1, 2, 3, 12, 34, 100
  print(ens & ens2)                             # 2, 3
  print(ens - ens2)                             # 1, 12
  for e in ens: print(e, end=" ")               # 1 2 3 12
  print()
```

Le module abc

- Le module *abc* (*Abstract Base Classes*) permet de définir des classes abstraites.
- Ce module fournit notamment la méta-classe *ABCMeta* qui permet de définir des classes abstraites.
- Il fournit également le décorateur *abstractmethod* pour définir des méthodes abstraites dans une classe dont la méta-classe est *ABCMeta* : la classe ne pourra être instanciée que si elle redéfinit toutes ses méthodes abstraites.
- Ce décorateur permet également de définir des méthodes de classes abstraites (on le combine avec *classmethod*) et des propriétés abstraites (on le combine avec *property*).
- Attention, en Python, le terme de *classe abstraite* n'est donc pas le même qu'en C++/Java/C# puisqu'une classe abstraite peut très bien ne pas avoir de méthode abstraites... (voir le premier exemple).
- Python fournit également les modules *numbers* et *collections* qui fournissent des classes abstraites pour les nombres et les collections (une classe qui hérite de *collections.Sequence* doit implémenter les méthodes *__getitem__* et *__len__*, par exemple, mais disposera automatiquement d'autres méthodes "mixins").

Exemple d'utilisation du module abc

```
from abc import ABCMeta, abstractmethod

class Figure(metaclass=ABCMeta):
  def __init__(self, nom):
    self._nom = nom

class FigureFermee(Figure):
  def __init__(self, nom):
    Figure.__init__(self, nom)

  @abstractmethod
  def surface(self):
    ...

class Rectangle(FigureFermee):
  def __init__(self, nom, coords, largeur, hauteur):
    FigureFermee.__init__(self, nom)
    self._coin_inf_gauche = coords
    self._largeur, self._hauteur = largeur, hauteur

  def surface(self):
    return self._largeur * self._hauteur

fig = Figure("abstraite")                          # Ok car pas de abstractmethod
fig_fermee = FigureFermee("toujours abstraite")    # TypeError: Can't instantiate abstract
                                                   # class FigureFermee with abstract methods surface
rect = Rectangle("carré", (1, 1), 10, 10)
rect.surface()                                     # 100
```

Exemple amélioré

```
class Figure(metaclass=ABCMeta):
  @abstractmethod
  def __init__(self, nom):
    self._nom = nom

  @property
  def nom(self):
    return self._nom

class FigureFermee(Figure):
  def __init__(self, nom):
    Figure.__init__(self, nom)

  @property
  @abstractmethod
  def surface(self):            # Toute figure fermée a une surface...
    ...

class Rectangle(FigureFermee):
  def __init__(self, nom, coords, largeur, hauteur):
    FigureFermee.__init__(self, nom)
    self._coin_inf_gauche = coords
    self._largeur, self._hauteur = largeur, hauteur

  @property
  def surface(self):          # Redéfinition de la propriété abstraite
    return self._largeur * self._hauteur

fig = Figure("abstraite")                          # TypeError: Can't instantiate abstract class...
rect = Rectangle("carré", (1, 1), 10, 10)
rect.surface                                       # 100
rect.nom                                           # 'carré'
```


Sérialisation et désérialisation

- La sérialisation consiste à transformer une structure de données en une suite d'octets afin de la stocker sur disque ou de l'envoyer sur le réseau. La désérialisation est l'opération inverse.
- Python dispose de plusieurs modules de sérialisation. Le plus utilisé est *pickle*, qui sérialise dans un format binaire uniquement lisible par Python.
- On peut également installer le module *yaml* qui sauvegarde les données dans un format lisible proche de XML ou le module *json* qui sauvegarde au format JSON (*JavaScript Object Notation*).
- Avec *pickle* et *json*, la sérialisation consiste à appeler les méthodes *dump* (pour écrire dans un fichier) ou *dumps* (pour écrire une chaîne d'octets) et la désérialisation consiste à appeler les méthodes *load* ou *loads*.
- Avec *yaml*, les méthodes *dump* et *load* traitent à la fois les fichiers et les chaînes d'octets.

Exemples

- Avec *pickle* :

```
import pickle

e1 = Ensemble(...)

# Sérialisation de e1 dans un fichier
with open('ensemble.pickle', 'wb') as f:
   pickle.dump(e1,f)                    # ou : pickle.dump(e1, open('ensemble.pickle', 'wb'))

# Désérialisation
with open('ensemble.pickle', 'rb') as f:
   e1 = pickle.load(f)                  # ou : e1 = pickle.load(open('ensemble.pickle', 'rb'))
```

- Avec YAML (il faut installer *PyYAML*) :

```
import yaml

e1 = Ensemble(...)

# Sérialisation de e1 dans un fichier
with open('ensemble.yaml', 'w') as f:
   yaml.dump(e1,f)                   # ou yaml.dump(e1, open('ensemble.yaml', 'w'))

# Désérialisation
with open('ensemble.yaml', 'r') as f:
   e1 = yaml.load(f)                 # ou e1 = yaml.load( open('ensemble.yaml', 'r'))
```


Sérialisation d'un Ensemble

```
import pickle, yaml
(...)
s = pickle.dumps(ens)                       # Sérialise dans un tableau d'octets
print(s)                                    # b'\x80\x03censemble\nEnsemble\nq\x00)...
ens3 = pickle.loads(s)
print(ens3)                                 # 1, 2, 3, 12

with open('ensemble.pickle', 'wb') as f:
  pickle.dump(ens, f)                       # Sérialise dans un fichier binaire

with open('ensemble.pickle', 'rb') as f:
  ens3 = pickle.load(f)

s = yaml.dump(ens)                          # YAML n'a qu'une méthode dump
print(s)                                    # !!python/object:ensemble.Ensemble
                                            # _Ensemble__corps: {1: true, 2: true, 3: true, 12: true}
ens4 = yaml.load(s)
print(ens3)                                 # 1, 2, 3, 12

with open('ensemble.yaml', 'w') as f:
  yaml.dump(ens, f)                       # Sérialise dans un fichier texte

with open('ensemble.yaml', 'r') as f:
  ens4 = yaml.load(f)
```

Fichiers texte

- La fonction *open()* permet d'ouvrir un fichier en lecture (mode *"r"* par défaut), en écriture (mode *"w"*) ou en ajout (mode *"a"*).
- Le mode peut être suivi de la lettre *b* pour indiquer une ouverture en mode binaire.
- En règle générale, il est préférable de considérer le fichier ouvert comme un itérateur classique (on peut alors utiliser une instruction *for*).
- Les méthodes *read()* et *readline()* renvoient une *chaîne* classique. La méthode *readlines()* renvoie un *tableau de chaînes* et n'ajoute pas de retour à la ligne.
- En lecture, la fin de fichier est détectée lorsque *read()* ou *readline()* renvoient une chaîne vide.
- La méthode *write()* écrit une chaîne dans le fichier ouvert en écriture. Elle renvoie le nombre de caractères écrits.
- On ferme un fichier avec la méthode *close()*.
- Si l'on a utilisé la construction *with*, le fichier ouvert est automatiquement fermé à la sortie du bloc.

Fichiers texte

Exemples :

```
fd = open("monfichier.txt")     # Ouverture en lecture par défaut
for ligne in fd:
   # faire quelque chose avec ligne
fd.close()                      # Fermeture explicite

for ligne in open("autre.txt"):
   # faire quelque chose avec ligne
# Ici le fichier est implicitement fermé car on est sorti de sa portée

fd = open("monfichier.txt")
contenu_complet = fd.read()     # une chaîne contenant tout le fichier

fd.seek(0)                      # on revient au début
prem = fd.readline()            # une chaîne contenant la 1ere ligne
sec  = fd.readline()            # une chaîne contenant la 2e ligne

fd.seek(0)
tab = fd.readlines()            # un tableau de lignes
print(tab[0])                   # première ligne...
```

Fichiers texte

Exemples :

```
fd = open("monfichier.txt", "w")     # Ouverture en écriture

# Si monfichier.txt existait, il a donc été écrasé !

fd.write("coucou\n")
fd.close()

with open("monfichier.txt") as fd:   # utilisation d'un bloc with
   # ici, fd est ouvert en lecture
# Sortie du bloc : fd est automatiquement fermé...

with open("monfichier.txt", "a") as fd:
  fd.write("au revoir\n")            # Ajout d'une ligne à la fin du fichier

from urllib.request import urlopen   # Gestion des URL comme des fichiers
import codecs

for ligne in urlopen("http://www.monsite.fr"):
   # on récupère des octets... donc il faut les décoder pour les transformer en caractères UTF-8
   print(codecs.decode(ligne))
```


Expressions régulières

Expressions régulières

- Les expressions régulières sont gérées via le module *re*, qu'il faut donc importer.
- Bien qu'elle ne soit pas nécessaire, la méthode *regexp = re.compile(motif)* permet d'optimiser les recherches. Elle permet également de passer des options (*re.I* ou *re.IGNORECASE*, par exemple).
- Si le *motif* contient des caractères spéciaux, on peut utiliser une *chaîne brute* en la préfixant du caractère *r* : *r"ceci est une chaîne brute"* (fonctionne également avec *'*, *'''* et *"""*).
- La méthode *search* d'une expression régulière renvoie un objet « match » si une correspondance a été trouvée, *None* sinon.
- La méthode *match* renvoie un objet « match » si la chaîne passée en paramètre correspond exactement au motif, *None* sinon.
- La méthode *sub* renvoie une chaîne où la première occurrence du motif dans la chaîne initiale a été remplacée par une autre chaîne.
- La méthode *findall* renvoie la liste de toutes les chaînes correspondant au motif dans la chaîne. La méthode *finditer* fait de même, mais renvoie un itérateur. On peut paramètrer le début et la fin de la recherche dans la chaîne.
- Voir les documentations complètes de ces méthodes dans la documentation du module *re*...

Opérations sur un objet « match »

Un objet « match » renvoyé par les appels à *search* ou *match* dispose des méthodes suivantes :

- *group(grp)* renvoie une chaîne correspond au texte capturé par l'expression ou la chaîne capturée par le groupe passé en paramètre (le groupe 0 correspond à toute la capture). Si l'on a utilisé des groupes nommés, on peut passer les noms en paramètre.
- *groups* renvoie un tuple contenant toutes les groupes capturés, à partir du groupe 1.
- *groupdict* renvoie un dictionnaire de tous les groupes nommés qui ont été capturés. Les clés sont les noms des groupes.
- *start* et *end* renvoient l'indice de début et de fin de la capture.

Rappels :

- Un groupe est délimité entre parenthèses. Un groupe est nommé par *?P<nom>*, *(?P<groupe>\d+)*, par exemple.
- Le traitement d'un groupe nommé est plus lent que celui d'un groupe non nommé.
- Ce qui a été capturé par un groupe est représenté par *\i* pour le groupe *i* ou par *(?P=nom)* pour le groupe nommé *nom*.
- Si l'on ne veut pas mémoriser le groupe capturé, on utilise la notation `(?: ...)`.

Exemple

On veut déterminer si un nom de fichier entré au clavier est un nom de fichier DOS valide, sachant que :

- Les noms de fichiers DOS ne sont pas sensibles à la casse
- Ils sont de la forme 8.3 : un nom principal et une extension, qui est facultative.
- Le nom principal doit commencer par une lettre et contenir des lettres des chiffres ou des blancs soulignés. Les lettres accentuées ne sont pas reconnues.
- L'extension ne contient que des lettres ou des chiffres.

Quelle expression régulière permettra de déterminer si la saisie est correcte et permettra d'isoler les deux parties du nom (nom principal et extension) ?

- Nom principal : `[a-zA-Z][a-zA-Z0-9_]{0,7}`
- Extension : `[a-zA-Z0-9]{1,3}`
- Nom complet, avec trois groupes pour capturer le nom principal et l'extension facultative : `([a-zA-Z][a-zA-Z0-9_]{0,7})(\.([a-zA-Z0-9]{1,3}))?`
- Idem, mais avec groupes nommés : `(?P<nom>[a-zA-Z][a-zA-Z0-9_]{1,7})(\.(?P<ext>[a-zA-Z0-9]{1,3}))?`

Exemple

```
import re

nom_dos = re.compile(r"(?P<nom>[a-zA-Z][a-zA-Z0-9_]{1,7})(\.(?P<ext>[a-zA-Z0-9]{1,3}))?")
nom_fic = "blabla.txt"                  # ou nom_fic = input("Nom du fichier : ")
result = re.match(nom_dos, nom_fic)
if result is not None:
    print(result.groups())          # ('blabla', '.txt', 'txt')
    nom, extension = result.group("nom"), result.group("ext")
    print(nom, extension)           # blabla txt
    nom, extension = result.group(1), result.group(3)
    print(nom, extension)           # blabla txt
else:
    print(nom_fic, "n'est pas un nom de fichier DOS")

nouveau_nom = re.sub(nom_dos, r"blibli\2", nom_fic)
print(nouveau_nom)                  # blibli.txt

print(re.sub(r"(\w+)\s+(\w+)", r"\2 \1", "mots inversés"))    # inversés mots

while True:
  age = input("Entrez un âge : ")
  if re.match("\d+", age) is not None:    # Plus besoin de try: ...
     age = int(age)                       # Mais age pourrait être une valeur non correcte...
     break
```