

# Python 103



Programmation concurrente (3.6.4)



# Histoire de la concurrence

- La concurrence est en fait dérivée des premiers travaux sur les chemins de fer et la télégraphie, d'où l'utilisation de noms tels que le sémaphore
- Essentiellement, il était nécessaire de gérer plusieurs trains sur le même réseau ferroviaire de manière à ce que chaque train puisse se rendre à destination en toute sécurité sans encourir d'accident
- Ce n'est que dans les années 1960 que les chercheurs se sont intéressés à l'informatique concurrente, et c'est Edsger W. Dijkstra qui a publié le premier article dans ce domaine, où il a identifié et résolu le problème de l'exclusion mutuelle
- Ensuite, Dijkstra a défini des concepts fondamentaux de concurrence, tels que les sémaphores, les exclusions mutuelles et les blocages, ainsi que l'algorithme du plus court chemin de Dijkstra

# Histoire de la concurrence

- La concurrence, comme dans la plupart des domaines de l'informatique, est encore un domaine incroyablement jeune comparé à d'autres domaines d'étude tels que les mathématiques, et cela vaut la peine de garder cela à l'esprit
- Il y a encore un énorme potentiel de changement dans le domaine, et il reste un domaine passionnant pour tous (universitaires, concepteurs de langage et développeurs)
- L'introduction de primitives de concurrence de haut niveau et une meilleure prise en charge du langage natif ont vraiment amélioré la façon dont nous, en tant qu'architectes logiciels, mettons en œuvre des solutions concurrentes
- Pendant des années, c'était incroyablement difficile à faire, mais avec l'avènement de nouvelles API concurrentes, et la maturation des framework et des langages, cela commence à devenir beaucoup plus facile pour nous en tant que développeurs

# Histoire de la concurrence

- Les concepteurs de langages sont confrontés à un défi important lorsqu'ils tentent d'implémenter une concurrence non seulement sûre, mais efficace et facile à écrire pour les utilisateurs de cette langue
- Les langages de programmation tels que Golang, Rust et même Python ont fait de grands progrès dans ce domaine, et il est beaucoup plus facile d'exploiter tout le potentiel des machines sur lesquelles vos programmes fonctionnent

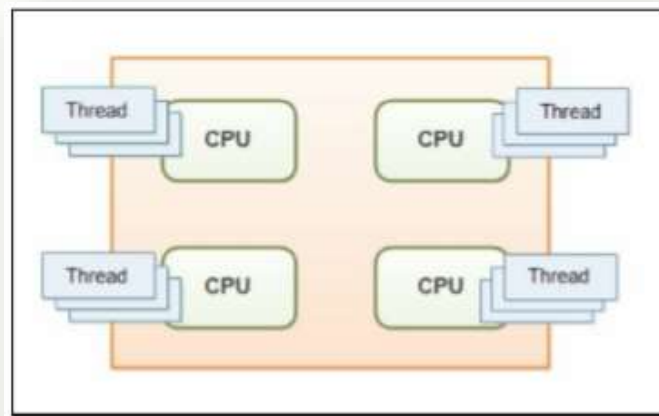
# Qu'est-ce qu'un thread ?

- Un thread peut être défini comme un flux ordonné d'instructions pouvant être programmé pour s'exécuter comme tel par les systèmes d'exploitation
- Ces threads, typiquement, vivent dans des processus, et consistent en un compteur de programme, une pile, et un ensemble de registres ainsi qu'un identifiant
- Ces threads sont la plus petite unité d'exécution à laquelle un processeur peut allouer du temps
- Les threads peuvent interagir avec des ressources partagées et la communication est possible entre plusieurs threads
- Ils sont également capables de partager de la mémoire, de lire et d'écrire différentes adresses de mémoire, mais c'est là un problème
- Lorsque deux threads commencent à partager de la mémoire et que vous n'avez aucun moyen de garantir l'ordre d'exécution d'un thread, vous pouvez commencer à voir des problèmes ou des bogues mineurs qui vous donnent des valeurs erronées ou font planter votre système



# Qu'est-ce qu'un thread ?

- La figure suivante montre comment plusieurs threads peuvent exister sur plusieurs processeurs différents



# Types de threads

- Dans un système d'exploitation typique, nous avons généralement deux types de threads distincts:
  - Threads au niveau utilisateur : Threads que nous pouvons activement créer, exécuter et tuer pour toutes nos tâches
  - Threads au niveau du noyau : Threads de très bas niveau au nom du système d'exploitation
- Python fonctionne au niveau de l'utilisateur, et donc, tout ce que nous couvrons dans ce document sera, principalement, axé sur ces threads au niveau de l'utilisateur

# Qu'est-ce que le multithreading ?

- Lorsque les gens parlent de processeurs multithread, ils font généralement référence à un processeur qui peut exécuter plusieurs threads simultanément, ce qu'ils sont capables de faire en utilisant un seul cœur capable de changer très rapidement le contexte entre plusieurs threads
- Ce contexte de commutation se déroule en si peu de temps que l'on pourrait penser que plusieurs threads fonctionnent en parallèle alors qu'en réalité ce n'est pas le cas
- Lorsque vous essayez de comprendre le multithreading, il est préférable de considérer un programme multithread comme un bureau
- Dans un programme monothread, il n'y aurait qu'une seule personne travaillant dans ce bureau à tout moment, manipulant tout le travail de manière séquentielle
- Cela deviendrait un problème si l'on considère ce qui se passe lorsque ce travailleur solitaire s'embourbe dans la paperasserie administrative et est incapable de passer à un travail différent



# Qu'est-ce que le multithreading ?

- Voyons quelques avantages du threading :
  - Les threads multiples sont excellents pour réduire le temps blocage des E/S liées aux programmes
  - Les threads sont légers en termes d'empreinte mémoire par rapport aux processus
  - Les threads partagent des ressources, et donc la communication entre eux est plus facile

# Qu'est-ce que le multithreading ?

- Il y a aussi des inconvénients, qui sont les suivants :
  - Les threads CPython sont entravés par les limitations du verrou global de l'interpréteur (GIL), dont nous parlerons plus en détail
  - Bien que la communication entre les threads puisse être plus facile, vous devez veiller à ne pas implémenter de code soumis à des conditions de concurrence
  - Il est coûteux en temps de changer de contexte entre plusieurs threads
  - En ajoutant plusieurs threads, vous pourriez voir une dégradation des performances globales de votre programme

# Qu'est-ce qu'un processus ?

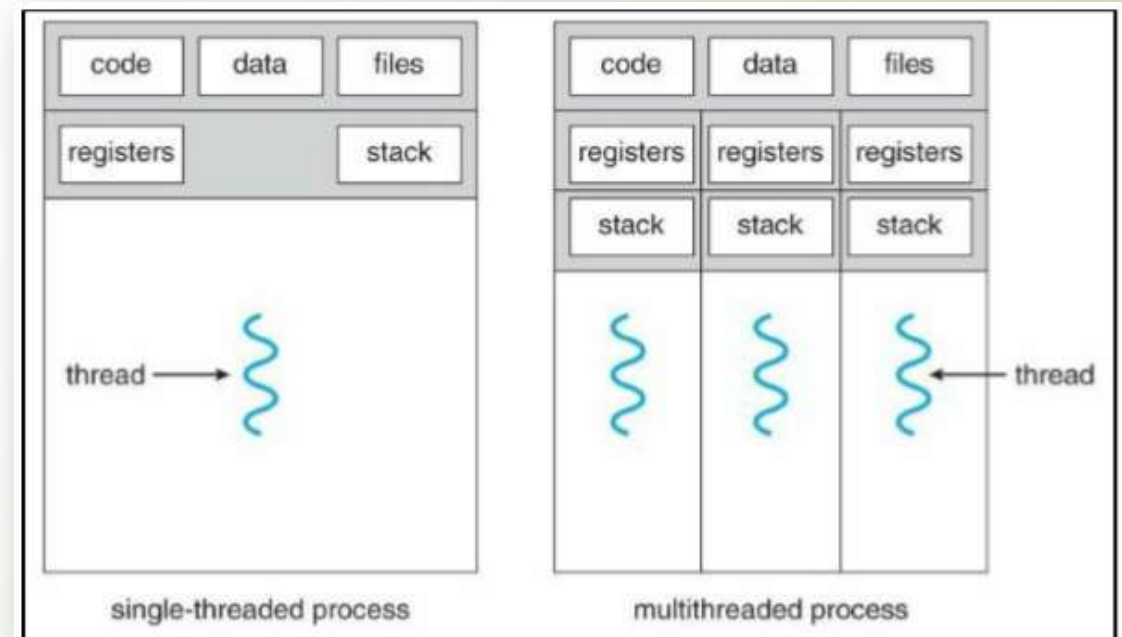
- Les processus sont très similaires aux threads
- Ils nous permettent de faire à peu près tout ce qu'un thread peut faire mais l'avantage principal est qu'ils ne sont pas liés à un noyau CPU singulier
- Si nous étendons notre analogie de bureau, cela signifie essentiellement que si nous avons un processeur à quatre cœurs, nous pourrions embaucher deux membres de l'équipe de vente dédiés et deux employés, et tous les quatre seraient en mesure d'exécuter le travail en parallèle
- Les processus sont également capables de travailler sur plusieurs choses à la fois, tout comme notre employé de bureau unique multithread

# Qu'est-ce qu'un processus ?

- Ces processus contiennent un thread principal primaire, mais peuvent engendrer plusieurs sous-threads qui contiennent chacun leur propre jeu de registres et une pile
- Ils peuvent devenir multithread si vous le souhaitez
- Tous les processus fournissent toutes les ressources dont l'ordinateur a besoin pour exécuter un programme

# Qu'est-ce qu'un processus ?

- Dans l'image suivante, vous voyez deux diagrammes côte-à-côte; les deux sont des exemples d'un processus
- Vous remarquez que le processus sur la gauche contient un seul thread, autrement connu comme le thread primaire
- Le processus sur la droite contient plusieurs threads, chacun avec son propre ensemble de registres et de piles





# Qu'est-ce qu'un processus ?

- Avec les processus, nous pouvons améliorer la vitesse de nos programmes dans des scénarios spécifiques où nos programmes sont liés au processeur et nécessitent plus de puissance CPU
- Cependant, en engendrant de multiples processus, nous sommes confrontés à de nouveaux défis en termes de communication interprocessus, et en tentant de ne pas entraver les performances en passant trop de temps sur cette communication interprocessus (IPC)

# Propriétés des processus ?

- Les processus UNIX sont créés par le système d'exploitation et contiennent généralement les éléments suivants
  - ID de processus
  - ID de groupe de processus
  - ID utilisateur et ID de groupe
  - Environnement
  - Répertoire de travail
  - Instructions de programme
  - Registres
  - Pile
  - Heap
  - Descripteurs de fichier
  - Actions de signal
  - Bibliothèques partagées
  - Outils de communication interprocessus (tels que files d'attente de messages, pipes, sémaphores ou mémoire partagée)

# Propriétés des processus ?

- Les avantages des processus sont les suivants :
  - Les processus peuvent mieux utiliser les processeurs multi cœurs
  - Ils sont meilleurs que les threads multiples pour gérer les tâches gourmandes en ressources
  - Nous pouvons contourner les limitations du GIL en engendrant plusieurs processus. Programme
  - Un processus qui se termine en erreur ne provoquera pas l'arrêt du programme entier

# Propriétés des processus ?

- Voici les principaux inconvénients des processus :
  - Pas de ressources partagées entre les processus
  - Nous devons mettre en œuvre une forme de IPC
  - Tout ceci nécessite plus de mémoire

# Multi processing ?

- En Python, nous pouvons choisir d'exécuter notre code à l'aide de plusieurs threads ou de plusieurs processus si nous souhaitons essayer d'améliorer les performances par rapport à une approche monothread standard
- Nous pouvons aller vers une approche multithread et être limité à la puissance de traitement d'un noyau de CPU, ou inversement,
- Nous pouvons aller avec une approche multi processing et utiliser le nombre total de cœurs de processeurs disponibles sur notre machine
- Dans les ordinateurs modernes d'aujourd'hui, nous avons tendance à avoir de nombreux processeurs et cœurs, de sorte que nous limiter à un seul, rend le reste de notre machine inactive
- Notre objectif est d'essayer d'extraire tout le potentiel de notre matériel et de nous assurer d'obtenir le meilleur rapport qualité-prix et de résoudre nos problèmes plus rapidement que quiconque



# Multi processing ?

- Avec le module multi processing de Python, nous pouvons utiliser efficacement le nombre total de cœurs et de processeurs, ce qui peut nous aider à obtenir de meilleures performances en ce qui concerne les problèmes liés au processeur



- La figure précédente montre un exemple de la façon dont un cœur CPU commence à déléguer des tâches à d'autres cœurs

# Multi processing ?

- Dans toutes les versions de Python inférieures ou égales à 2.6, nous pouvons obtenir le nombre de cœurs de processeur disponibles en utilisant le code suivant

```
>>> import multiprocessing
>>> multiprocessing.cpu_count()
8
>>>
```

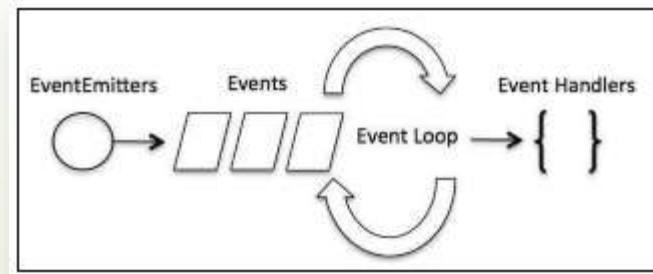
- Non seulement le multi processing nous permet d'utiliser plus de notre machine, mais nous évitons également les limitations que le verrou global d'interpréteur nous impose dans Python
- Un inconvénient potentiel de plusieurs processus est que nous n'avons intrinsèquement pas d'état partagé et que nous manquons de communication
- Nous devons donc passer à travers une forme de IPC, et la performance peut en prendre un coup
- Cependant, ce manque d'état partagé peut faciliter le travail, car vous n'avez pas à vous battre contre des conditions de concurrence potentielles dans votre code

# Programmation événementielle

- La programmation événementielle est une partie importante de nos vies
- Nous en voyons des exemples tous les jours lorsque nous ouvrons notre téléphone ou travaillons sur notre ordinateur
- Ces appareils fonctionnent uniquement de manière événementielle
- Par exemple, lorsque vous cliquez sur une icône sur votre bureau, le système d'exploitation enregistre cela comme un événement, puis effectue l'action nécessaire liée à ce style d'événement spécifique.

# Programmation événementielle

- Chaque interaction que nous faisons peut être caractérisée comme un événement ou une série d'événements, et ceux-ci déclenchent généralement des rappels (callbacks)
- Si vous avez une expérience antérieure avec JavaScript, vous devriez vous familiariser avec ce concept de rappel et le modèle de conception de rappel
- En JavaScript, le cas d'utilisation prédominant pour les rappels est lorsque vous effectuez des requêtes HTTP RESTful et que vous voulez pouvoir effectuer une action lorsque vous savez que cette action s'est terminée avec succès et que nous avons reçu notre réponse HTTP:



# Programmation événementielle

- Si nous regardons l'image précédente, elle nous montre un exemple de la façon dont les programmes pilotés par les événements traitent les événements
- Nous avons nos EventEmitters sur le côté gauche
- Ces derniers déclenchent plusieurs événements, qui sont captés par la boucle d'événements de notre programme, et, s'ils correspondent à un gestionnaire d'événements prédéfini, ce gestionnaire est alors déclenché pour gérer ledit événement



# Programmation événementielle

- Les rappels (callbacks) sont souvent utilisés dans les scénarios où une action est asynchrone
- Supposons, par exemple, que vous postuliez pour un emploi chez Google, que vous leur donniez une adresse e-mail, et qu'ils vous contacteront quand ils prendront leur décision
- Cela revient essentiellement à enregistrer un rappel sauf qu'au lieu de vous envoyer un e-mail, vous exécutez un code arbitraire chaque fois que le rappel est appelé

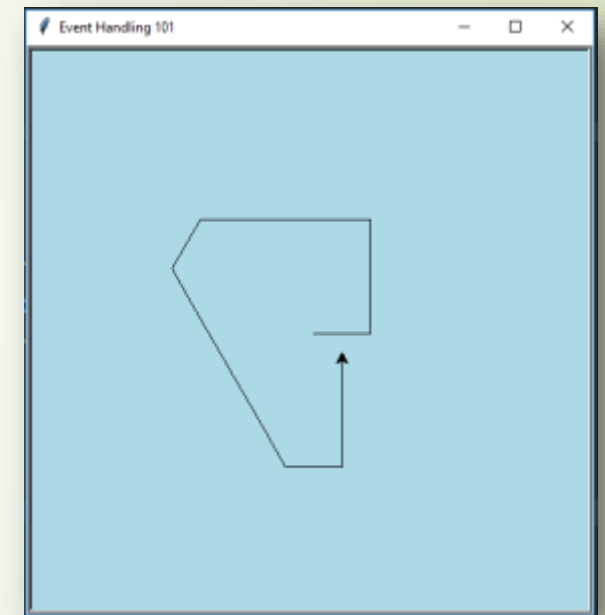
# Turtle

- Turtle est un module graphique qui a été écrit en Python, et c'est un excellent point de départ pour intéresser les enfants à la programmation
- Il gère toutes les complexités liées à la programmation graphique et leur permet de se concentrer uniquement sur l'apprentissage des bases tout en les gardant intéressés
- C'est aussi un très bon outil à utiliser pour démontrer des programmes axés sur les événements
- Il comporte des gestionnaires d'événements et des auditeurs, ce qui est tout ce dont nous avons besoin

```
turtle1.py x
1  #!/usr/bin/env python
2  # -*- coding: utf-8 -*-
3
4  import turtle
5
6
7  def main():
8      import turtle
9      turtle.setup(500, 500)
10     window = turtle.Screen()
11     window.title("Event Handling 101")
12     window.bgcolor("lightblue")
13     nathan = turtle.Turtle()
14
15     def moveForward():
16         nathan.forward(50)
17
18     def moveLeft():
19         nathan.left(30)
20
21     def moveRight():
22         nathan.right(30)
23
24     def start():
25         window.onkey(moveForward, "Up")
26
27     window.onkey(moveForward, "Up")
28     window.onkey(moveLeft, "Left")
29     window.onkey(moveRight, "Right")
30     window.listen()
31     window.mainloop()
32
33
34  if __name__ == '__main__':
35     main()
36
```

# Turtle

- Dans la première ligne de cet exemple de code précédent, nous importons le module graphique Turtle
- Nous poursuivons ensuite pour mettre en place une fenêtre de tortue de base avec le titre Event Handling 101 et une couleur de fond de bleu clair
- Une fois la configuration initiale terminée, nous définissons trois gestionnaires d'événements distincts :
- `moveForward` : Quand nous voulons déplacer notre personnage de 50 unités
- `moveLeft/moveRight` : Quand nous voulons faire pivoter notre personnage dans les deux sens de 30 degrés
- Une fois que nous avons défini nos trois gestionnaires distincts, nous passons ensuite à la mise en correspondance de ces gestionnaires d'événements avec les touches vers le haut, vers la gauche et vers la droite en utilisant la méthode `onkey`



# Programmation réactive

- La programmation réactive est très similaire à celle de l'événement, mais au lieu de tourner autour des événements, elle se concentre sur les données
- Plus spécifiquement, elle traite des flux de données et réagit aux changements de données spécifiques

# ReactiveX - RxPy

- RxPy est l'équivalent Python du très populaire framework ReactiveX
- Si vous avez déjà effectué une programmation dans Angular 2 et versions suivantes, vous l'utiliserez lors de l'interaction avec les services HTTP
- Ce framework est une conglomération du modèle d'observateur, du modèle d'itérateur et de la programmation fonctionnelle
- Nous nous abonnons essentiellement à différents flux de données entrantes, puis créons des observateurs qui écoutent les événements spécifiques déclenchés
- Lorsque ces observateurs sont déclenchés, ils exécutent le code correspondant à ce qui vient de se passer



# ReactiveX - RxPy

- Nous considérerons un centre de données comme un bon exemple de la façon dont la programmation réactive peut être utilisée
- Imaginez que ce centre de données possède des milliers de racks de serveurs, qui calculent en permanence des millions et des millions de calculs
- L'un des plus grands défis de ces datacenters est de garder suffisamment refroidis tous ces racks de serveur bien serrés pour qu'ils ne s'endommagent pas
- Nous pourrions installer plusieurs thermomètres dans notre centre de données pour nous assurer de ne pas avoir trop chaud et envoyer les lectures de ces thermomètres à un ordinateur central en continu

# ReactiveX - RxPy

- Au sein de notre centrale de contrôle, nous pourrions mettre en place un programme RxPy qui observe ce flux continu d'informations sur la température
- Au sein de ces observateurs, nous pourrions alors définir une série d'événements conditionnels à écouter, puis réagir chaque fois que l'un de ces conditionnels est atteint
- Un tel exemple serait un événement qui se déclenche uniquement si la température d'une partie spécifique du centre de données devient trop chaude
- Lorsque cet événement est déclenché, nous pouvons alors automatiquement réagir et augmenter le débit de tout système de refroidissement vers cette zone particulière, et ainsi ramener la température à un bon niveau

# Programmation GPU

- Les GPU sont réputés pour leur capacité à rendre des jeux vidéo d'action rapides et de haute résolution
- Ils sont capables d'absorber ensemble les millions de calculs nécessaires par seconde afin de s'assurer que chaque vertex des modèles 3D de votre jeu est au bon endroit, et qu'ils sont mis à jour toutes les quelques millisecondes afin d'assurer une douceur de 60 FPS
- De manière générale, les GPU sont incroyablement performants pour effectuer la même tâche en parallèle, des millions et des millions de fois par minute
- Mais si les GPU sont si performants, alors pourquoi ne pas les utiliser à la place de nos processeurs ?
- Alors que les GPU peuvent être incroyablement performants au traitement graphique, ils ne sont cependant pas conçus pour gérer les subtilités de l'exécution d'un système d'exploitation et de l'informatique à usage général
- Les CPU ont moins de noyaux, qui sont spécifiquement conçus pour la vitesse quand il s'agit de changer de contexte entre les tâches d'exploitation
- Si les GPU avaient les mêmes tâches, vous constateriez une dégradation considérable des performances globales de votre ordinateur

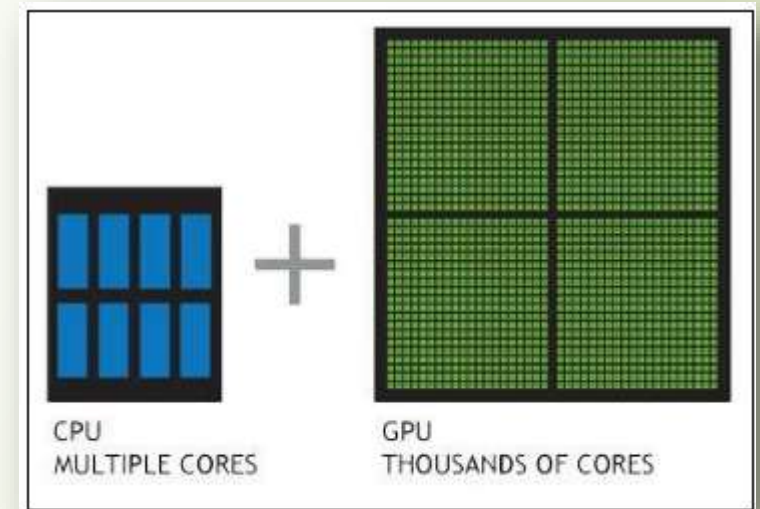
# Programmation GPU

- Mais comment pouvons-nous utiliser ces cartes graphiques de haute puissance pour autre chose que la programmation graphique ?
- C'est ici qu'interviennent les bibliothèques telles que PyCUDA, OpenCL et Theano
- Ces bibliothèques tentent de faire abstraction du code complexe de bas niveau auquel les API graphiques doivent interagir pour utiliser le GPU
- Ils simplifient grandement la réutilisation des milliers de cœurs de traitement plus petits disponibles sur le GPU et leur utilisation pour nos programmes coûteux en calcul



# Programmation GPU

- Ces unités de traitement graphique (GPU) encapsulent tout ce que les langages de script ne sont pas
- Ils sont hautement parallélisables et construits pour un débit maximum
- En les utilisant en Python, nous sommes en mesure d'obtenir le meilleur des deux mondes
- Nous pouvons utiliser un langage favorisé par des millions de personnes en raison de sa facilité d'utilisation, et rendre nos programmes incroyablement performants
- Les principales bibliothèques Python pour la programmation GPU : PyCUDA (spécifique Nvidia), OpenCL, Theano





# Références

- Python.org : <https://www.python.org/>
- Learning Python : <https://github.com/thierydecker/learning-python>
- ...

# Outils

- IDE Pycharm Community : <https://www.jetbrains.com/pycharm/>
- Analyse en ligne de code Python : <http://www.pythontutor.com/>
- ...