

Inleiding tot Python

Brecht Baeten

16 september 2015

1 Wat is Python?

Python is een open source programmeer taal. Voor ingenieurstoepassingen of wetenschappelijke projecten is Python interessant aangezien er een aantal Python modules bestaan die hier specifiek op gericht zijn. Deze modules maken het beheren van data of uitvoeren van complexe wiskundige algoritmes zeer eenvoudig, vandaar de populariteit. Python werkt volledig object georiënteerd, dit laat toe efficiënte en leesbare code te schrijven. Het schrijven van leesbare code is één van de hoekstenen van Python aangezien code veel vaker gelezen wordt dan geschreven. Python code hoeft ook niet gecompileerd te worden (althans niet door de gebruiker). Python is een geïnterpreteerde programmeertaal wat inhoudt dat code rechtstreeks uitgevoerd kan worden zonder eerst te compileren. Een ander groot voordeel van Python boven bijvoorbeeld Matlab is dat het gratis is! Iedereen, wetenschappers, studenten, particulieren, grote of kleine bedrijven kan Python gratis downloaden en gebruiken. In veel Linux distributies wordt Python standaard meegeleverd wat zelfs de installatie overbodig maakt.

2 Installatie in Windows

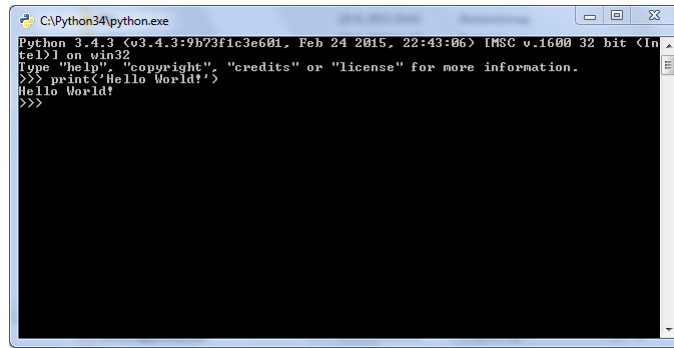
De nieuwste versie van Python kan worden gedownload via <https://www.python.org/downloads/>. Hier vind je de broncode en installer executables voor verschillende platformen. Downloaden, uitvoeren en klaar. Tijdens de installatie is het interessant om Python aan je Windows search path toe te voegen, zo kan je python steeds van in een command line interface openen. Het is ook interessant om de installatie locatie te veranderen naar bijvoorbeeld "C://python34".

Je kan nu al Python starten door op python.exe te klikken. Er opent een console waarin je python commando's kan invoeren (zie Figuur 1). Typ hier bijvoorbeeld:

```
print('Hello World!')
```

Druk op enter en je hebt je eerste python commando uitgevoerd.

Python code kan opgeslagen worden in eenvoudige tekst bestanden met de extensie ".py". Om deze te schrijven is een goede text editor met syntax highlighting aan te raden. In Windows is *notepad++* een goed, open source alternatief met syntax highlighting voor Python en een groot aantal andere talen ingebouwd. *notepad++* kan worden gedownload vanop <https://notepad-plus-plus.org/download/>.



Figuur 1: Python console met eerste commando

Omdat Python vanuit een commandline interface zal worden aangeroepen is het interessant om een deftige console te installeren. *ConEmu* is een open source console emulator met gelijkenissen aan *bash* in Linux. Deze is te downloaden vanaf <http://sourceforge.net/projects/conemu/files/latest/download>. Simpelweg downloaden, unzippen in een folder naar keuze en klaar.

3 Scripts

Omdat het telkens opnieuw invoeren van commando's achter elkaar niet zo praktisch is, is het interessant om een reeks commando's op te slaan in een script en dit uit te voeren. Open een tekst-editor en typ "**print**('Hello World!')" en sla het bestand op als `hello_world.py`. Om het bestand uit te voeren open je een console, navigeer naar de locatie waar je het bestand hebt opgeslagen, typ "`python hello_world.py`" en druk op enter.

Python heeft een aantal interessanten ingebouwde variabele types. Open een nieuw bestand, typ onderstaande commando's en sla het op als `variables.py`.

```
#variables
# integer
A = 1
print (A)

# float
B = 3.572
print (B)

# string
C = 'string'
print (C)

# list
D = [1,2,3]
print (D)

E = [1,'test',4,D]
print (E)
print (E[:2])
print (E[-1])

# dictionary
```

```

F = {3:1, 5:'test', 'spam':'eggs'}
print(F)
print(F['spam'])

# assignment by reference
G = F
G['spam'] = 'african swallow'
print(F)
print(G)

G = dict(F)
G['spam'] = 'european swallow'
print(F)
print(G)

```

Voor het script ditmaal uit met het commando "python -i variables.py". De "-i" zorgt ervoor dat python na het uitvoeren van alle commando's in het script naar de interactieve console gaat. Alle reeds gedefinieerde variabelen zijn nu ook beschikbaar in de console. Typ bijvoorbeeld:

```
E[3]
```

of:

```
A+B
```

Door "Ctrl+Z" in te geven kan je nu Python afsluiten en terugkeren naar de command line interface.

Zoals te merken in het voorgaande voorbeeld is Python zeer flexibel met data types. Zo kunnen een *Integer* en een *Float* met elkaar opgeteld worden en kan een *List* elementen met verschillende datatypes bevatten. Een zeer interessante data type is het *Dictionary* of *dict*. Hierin kunnen key / value paren worden opgeslagen en terug opgeroepen. Zowat elke Python variabele kan een *Dictionary* key zijn wat dit type zeer flexibel maakt. Het is interessant om even in te gaan op de manier van indexeren in Python. Een python *List* start bij index 0. Het laatste element van een *List* kan je oproepen met de index -1, het voorlaatste met index -2, enz. Je kan een deel van de lijst opvragen met behulp van de ":" operator, de subset loopt dan van de index voor de ":" tot (niet tot en met) de index erna. Indien er geen index voor of achter de ":" staat zal de subset starten of eindigen bij respectievelijk het eerste of het laatste element.

Er moet even aandacht besteed worden aan de manier waarop variabelen in Python gebruikt worden. In Python zijn variabelen slechts namen die naar een bepaalde waarde verwijzen. Wanneer een variabele toegewezen wordt, wordt enkel de referentie naar de waarde toegewezen. In het voorbeeld hierboven zullen "**print**(F)" en "**print**(G)" hetzelfde resultaat geven. Hierboven wordt "H" echter gedefinieerd als een nieuwe *Dictionary* met dezelfde waarden als "F". De commando's "**print**(F)" en "**print**(H)" zullen dus een verschillend resultaat hebben.

Echt programmeren begint pas bij maken van loops en voorwaarden: flow control. In python is dit zeer eenvoudig met behulp van "for" lussen of "if else" structuren:

```

# flow control
for i in range(10):
    if i%2 == 0:
        print('{} is even'.format(i))
    else:
        print('{} is oneven'.format(i))

```

In Python moet elk flow control element eindigen met een ":". De code binnen het element moet een tab inspringen. Dit zorgt voor een zekere leesbaarheid in de code. De "range(10)" functie in de for lus maakt een *List* met waarden van 0 tot en met 9. De ".format(i)" functie formatteert zijn inhoud op de plaats van de accolade's in de voorafgaande string. De format functie is eigenlijk een methode van het *String* datatype en geeft een hele reeks formatteer opties. Even zoeken op het net leert je heel veel over zulke ingebouwde functies.

Je kan over zowat elk datatype itereren. Een lus over een dictionary kan bijvoorbeeld op verschillende manieren. Ook itereren over verschillende variabelen tegelijk is eenvoudig met de "zip()" functie. Indien je itereert over een *List* en de index van de variabelen wil gebruiken kan je de functie "enumerate()" gebruiken:

```
# flow control advanced
A = {'foo': 'bar',
     'spam': 'eggs',
     'bacon': 'spam'}

# dictionaries
for key in A:
    print('{}: {}'.format(key,A[key]))

for key,val in A.iteritems():
    print('{}: {}'.format(key,val))

# zip
B = [1,2,3]
C = [7,8,9]
for b,c in zip(B,C):
    print('b = {:.0f}, c = {:.3f}'.format(b,c))

# enumerate
D = [1,11,111,1111]
for i,d in enumerate(D):
    print('D[{}] = {}'.format(i,d))
```

Een set commando's die vaak op dezelfde manier gebruikt worden kan je groeperen in een functie. Een functie definieer je met het "def" keyword, gevolgd door de functie naam, de argument namen tussen haakjes en een dubbelpunt. Argumenten kunnen verplicht of optioneel zijn. voor optionele argumenten moet een default waarde opgegeven worden in de functie definitie. Alle verplichte argumenten moeten ook voor de optionle komen in de functie definitie. Wanneer een functie veel optionele argumenten heeft is het gemakkelijk om deze via naam-waarde paren op te geven, de volgorde is dan niet meer van belang:

```
# functions
def spam(A,B,C=0,D=0):
    val = 100*A+10*B+C+0.1*D
    return val

D = spam(4,1)
print(D)

E = spam(4,1,5.3)
print(E)

F = spam(4,1,D=3,C=5)
print(F)
```

```
print(val)
```

Variabelen die binnen een functie gedefinieerd zijn bestaan ook enkel in de functie namespace. Bovenstaande code geeft dus een error omdat `val` enkel in de functie namespace gedefinieerd is, niet in de globale namespace. Functies kunnen wel gebruik maken van variabelen in de globale namespace maar niet omgekeerd. Wanneer de bovenstaande code runt krijg je dan ook de volgende foutmelding:

```
Traceback (most recent call last):
  File "C:\examples\functions.py", line 10, in <module>
    print(val)
NameError: name 'val' is not defined
```

4 NumPy, SciPy en Matplotlib

Er bestaan enorm veel modules waarin specifieke functie gedefinieerd zijn voor gebruik in Python. Twee voor Ingenieurs interessante modules zijn *NumPy*, een lineair algebra module, *SciPy*, een algemeen wetenschappelijke wiskunde module en *Matplotlib*, een module voor het maken van wetenschappelijke figuren van hoge kwaliteit. Windows installers kunnen worden gedownload via <http://sourceforge.net/projects/numpy/files/NumPy/>, <http://sourceforge.net/projects/scipy/files/scipy/> en <http://matplotlib.org/downloads.html>. Kies de binary kiezen die overeenkomt met jou systeem architectuur en python versie, downloaden en installeren. De meeste modules lopen steeds één versie achter op de laatste Python release, het kan dus interessant zijn om een iets oudere versie van Python te installeren.

Om een module te kunnen gebruiken in een script moet deze eerst geïmporteerd worden. Dit gebeurt met een **"import"** statement en kan op verschillende manieren. In Python blijft elke geïmporteerde module zijn eigen namespace behouden. Je kan dus de `linspace` functie uit *NumPy* aanroepen als `numpy.linspace(...)`. Om de notatie wat te verkorten kan je een module importeren onder een zelf gekozen naam met **"import ... as ..."** zoals aangegeven in het voorbeeld hieronder. Voor meer info over *NumPy* of *SciPy* wordt verwezen naar de online documentatie op <http://docs.scipy.org/doc/numpy/> en <http://docs.scipy.org/doc/scipy/reference/>

Matplotlib kan latex gebruiken om tekst weer te geven, dit kan permanent gedaan worden door de `'rc'` parameters aan te passen in het `matplotlibrc` bestand of deze in het script zelf aan te passen zoals hieronder aangegeven. De grootte van de figuur wordt default in inch opgegeven vandaar dat een correctiefactor naar centimeter. Met behulp van de `"savefig"` functie wordt de figuur opgeslagen als pdf (Figuur 2). Een zeer groot aantal voorbeelden van *Matplotlib* figuren en de code gebruikt om ze te maken is terug te vinden op <http://matplotlib.org/gallery.html>.

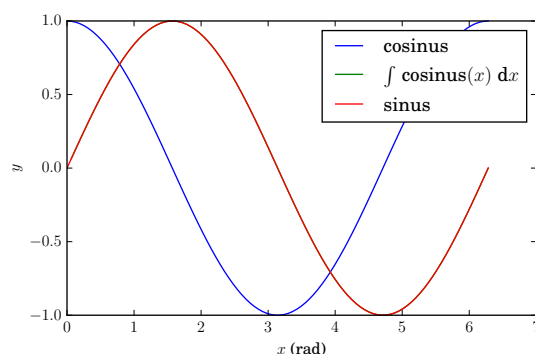
```
import numpy as np
import scipy.integrate
import matplotlib.pyplot as plt

# numpy example
x = np.linspace(0, 2*np.pi, 100)
c = np.cos(x)
s = np.sin(x)
ss = scipy.integrate.cumtrapz(c, x)
```

```
# matplotlib example
plt.rc('text', usetex=True)
plt.rc('font', family='serif')
plt.rc('figure', autolayout=True)

plt.figure(figsize=(15/2.54,10/2.54))
plt.plot(x,c,label='cosinus')
plt.plot(x[1:],ss,label='$\int \cosinus(x) dx$')
plt.plot(x,s,label='sinus')
plt.gca().set_xlabel('$x$ (rad)')
plt.gca().set_ylabel('$y$')
plt.legend()

plt.savefig('sinus_cosinus.pdf')
plt.show()
```



Figuur 2: Matplotlib voorbeeld

5 Object georiënteerd

Een grote sterkte van Python is de zeer grote Object gerichtheid van de programmeertaal. Zowat alles in python is een object met zijn methodes en kan aangepast of gebruikt worden in child classes. Een klasse wordt in python gedefinieerd met het keyword "**class**" gevolgd door de classenaam en een dubbelpunt. Opnieuw dient alles wat binnen de klasse definitie valt één tab verschoven te zijn tenopzichte van de klasse definitie. Binnen een klasse zijn er een aantal belangrijke methodes die kunnen gebruikt worden. De eerste is de "`__init__`" methode. Deze wordt gebruikt om een object instance te creëren. Het eerste argument is steeds "`self`", dit wordt gebruikt om attributen aan de instance toe te kennen, verder kunnen de argumenten op dezelfde wijze als bij functies gedefinieerd worden. De "`__`" voor en na de methode definitie geven aan dat deze een ingebouwde functie heeft. In het voorbeeld hieronder worden ook de "`__str__`" en "`__add__`" methodes gedefinieerd die respectievelijk het printen van een vector en de som van 2 vectoren definiëren. Telkens is "`self`" het eerste argument om een referentie naar de huidige object instance te hebben.

In het voorbeeld hieronder wordt ook een `plot` methode gedefinieerd die gebruikt kan worden om de vector te plotten. Onmiddellijk na de definitie volgt de documentatie van de functie tussen drie aanhalingstekens "`'''`". Deze wordt weergegeven indien de "`help()`" functie wordt opgeroepen, dit kan trouwens ook bij gewone functies. Het is een goede gewoonte om hier steeds te schrijven wat de functie doet, wat ze returned en wat de argumenten betekenen. Een andere goede gewoonte is

om methode's die bedoelt zijn om enkel binnen de klasse definitie te gebruiken te beginnen met een `"_"`. Op die manier is het voor een gebruiker van de klasse meteen duidelijk dat deze methode niet voor hem of haar bedoeld is.

In de code hieronder is het gedeelte dat rechtstreeks uitgevoerd moet worden in een `"if __name__ == '__main__':"` geplaatst. Dit zorgt ervoor dat deze code enkel uitgevoerd wordt als het script rechtstreeks wordt uitgevoerd. Meer hierover in de volgende sectie.

```
import matplotlib.pyplot as plt

# class definition
class vector:
    """
    A class defining a 2 dimensional vector
    """
    def __init__(self,x,y):
        """
        Returns a vector Instance

        Arguments:
        x: x coordinate of the vector
        y: y coordinate of the vector
        """
        self.x = x
        self.y = y

    def __str__(self):
        return '({},{})'.format(self.x,self.y)

    def __add__(self,v):
        return vector(self.x+v.x,self.y+v.y)

    def plot(self,origin=None):
        """
        Plots the vector from an optional origin

        Arguments:
        optional = None: starting point of the vector
        """
        if origin == None:
            origin = vector(0,0)

        plt.plot([origin.x,origin.x+self.x],[origin.y,origin.y+self.y])

if __name__ == '__main__':
    help(vector.plot)

    a = vector(1,4)
    b = vector(3,2)

    c = a+b

    print(a)
    print(b)
    print(c)
```

```

a.plot()
b.plot(a)
c.plot()
plt.show()

```

6 Eigen modules schrijven

Wanneer een project groeit komt er een moment dat het niet meer interessant is om alle code in één file te schrijven. In Python kan dit door het project op te delen in modules en deze in een hoofdbestand te importeren en aan te gebruiken. Een python module kan je definiëren door in een folder een bestand "`__init__.py`" aan te maken. In dit bestand kan je code zetten die uitgevoerd dient te worden tijdens de initialisatie van de module, maar je kan deze ook gewoon leeg laten. Wanneer er een "`__init__.py`" in een folder staat kan je de verschillende files in die folder importeren via het "`import`" commando zoals in het voorbeeld hieronder.

Wanneer je dit doet zullen alle commando's uit het geïmporteerde file uitgevoerd worden en de functie of klasse definities uit die file worden beschikbaar via de "." notatie. Nu wordt ook duidelijk waarom hierboven een deel van de code binnen een "`if __name__ == '__main__':`" constructie geplaatst is. Wanneer een script rechtstreeks uitgevoerd wordt zal Python een "`__name__`" veranderlijke aanmaken met de waarde "`'__main__'`". Wanneer een script geïmporteed wordt, wordt in de namespace van die module de veranderlijke "`__name__`" aangemaakt met de naam van de file als waarde. Dit kan interessant zijn bij het testen van submodules. Je kan steeds kijken welke variabelen er gedefinieerd zijn in de huidige namespace met de "`dir()`" functie.

In het voorbeeld hieronder wordt ook een submodule uit de subfolder `mymodule` aangeroepen. Deze bevat een functie die opnieuw via de "." notatie kan worden gebruikt. Via de "`dir(mymodule.submodule)`" functie kan je ook achterhalen welke functie allemaal gedefinieerd zijn in een bepaalde module. Wanneer een folder geïmporteed wordt worden enkel de commando's in het "`__init__.py`" script in die folder uitgevoerd. Je kan dan bijvoorbeeld in dit script een aantal andere scripts importeren die zo ook beschikbaar worden in het hoofd script.

```

import object_oriented
import mymodule.submodule

# what is in the current namespace
print( dir() )
print( dir(mymodule.submodule) )

# a file imported from the current directory
a = object_oriented.vector(8,2)
b = object_oriented.vector(4,6)
print(a+b)

# a file imported from a subdirectory
mymodule.submodule.test()

```

Python biedt ook de mogelijkheid om alle definities uit een bepaalde module te importeren in de huidige namespace. Dit gebeurt met het "`from mymodule.submodule import *`" commando. Eens dit is uitgevoerd zijn alle functies klassen en variabelen gedefinieerd in "`mymodule\submodule.py`" rechtstreeks beschikbaar, zonder "." notatie.

Door op deze manier een project goed te structureren blijft het geheel leesbaar en overzichtelijk.

© ⓘ 2015 Brecht Baeten

Dit werk is gelicenseerd onder de licentie Creative Commons Naamsvermelding-GelijkDelen 4.0 Internationaal. Ga naar <http://creativecommons.org/licenses/by-sa/4.0/> om een kopie van de licentie te kunnen lezen.