



Πανεπιστήμιο Κύπρου – Τμήμα Πληροφορικής
[ΕΠΛ371 – Προγραμματισμός Συστημάτων]

Overlap

ΟΜΑΔΑ 3

Αλεξάνδρου Σάββας

Ανδρέου Ανδρέας


Παπαφιλίππου Φίλιππος

Στεφανίδης Δημοσθένης

Εισαγωγή

Θέμα / Όνομα / Έργο



- Θέμα: Cycling Maps with  python™
- Όνομα: “OverLap”
- Έργο:

“Implement a Crawler for the CyclingMaps project. Particularly, your system has to provide a login box, through which a user provides login credentials and a favorite sport activity tracking website (e.g., runtastic + 2 more). You automatically download, align the trajectory data that is subsequently presented on a google map. The given data will be the foundation for an automated Cyprus Cycling Map we are trying to construct.”

PYTHON

Εισαγωγή

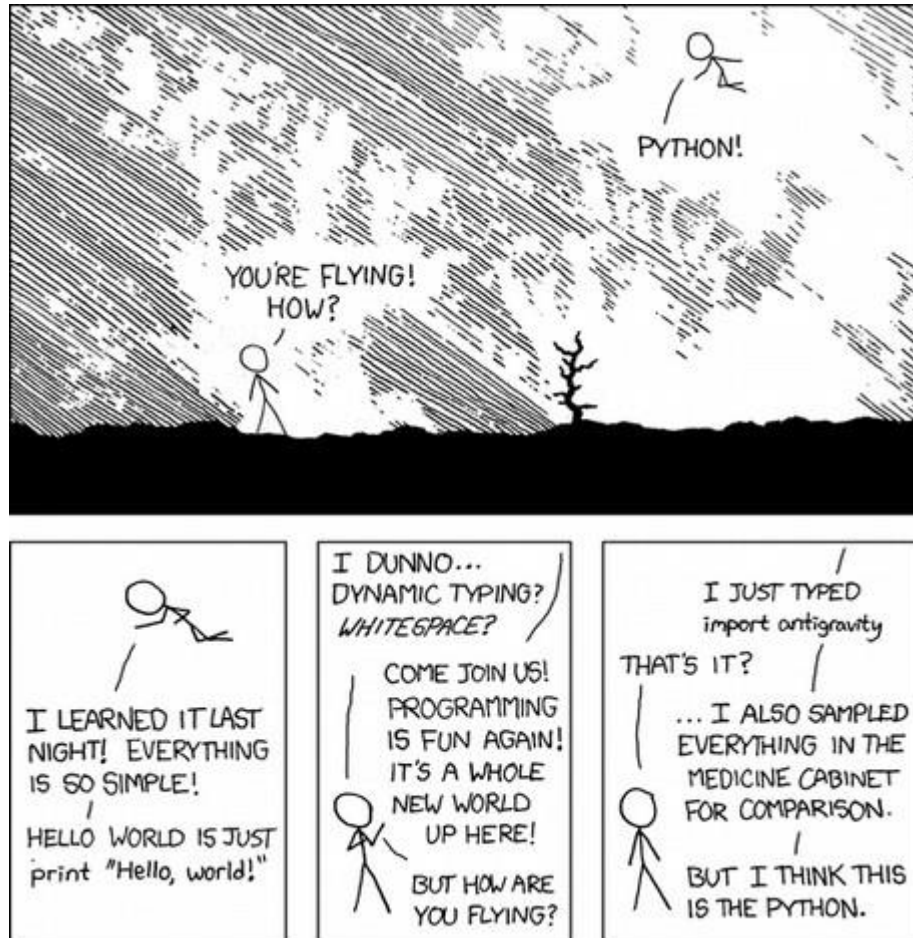
OverLap

Python

Λίγα Λόγια, Παραδείγματα

PYTHON

Λίγα λόγια



*I wrote 20 short programs in Python yesterday.
It was wonderful. Perl, I'm leaving you.*

PYTHON

Creator

- Ολλανδός
- 59 Χρονών
- Τίτλος: «Benevolent Dictator for Life».
- Βραβείο από το Free Software Foundation.
- Δουλεύει στη Dropbox.



Guido van Rossum

PYTHON

Ονομασία

OverLap



Monty Python's Flying Circus
[1969 – 1974]



Spam
[1970]

- Beautiful is better than ugly
- Explicit is better than implicit
- Simple is better than complex
- Complex is better than complicated
- Readability counts

~~TIMTOWTDIBSCINABTE~~



“There should be one— and preferably only one —obvious way to do it.”

- Google, Yahoo, CERN, NASA
- Χρησιμοποιείται σαν scripting language σε web applications – π.χ. μέσω του wsgi για Apache.
- Web app framework: Django, Pylons, Pyramid
- Βιβλιοθήκες: NumPy, SciPy, Matplotlib, BioPython, Astropy, Sage (για Μαθηματικά).
- Έχει ενσωματωθεί σε λογισμικά FreeCad, Blender, Maya, GIMP, Inkscape, Raspberry Pi.

PYTHON

Software

OverLap



- Εμπειρική μελέτη έδειξε πως οι scripting γλώσσες (όπως η python) είναι πιο παραγωγικές από τις συμβατικές γλώσσες (όπως Java, C) για προγράμματα που έχουν να κάνουν με string manipulation και αναζήτηση σε λεξικό. Η χρήση μνήμης έδειξε να ήταν συχνά καλύτερη από τη Java και όχι πολύ χειρότερη από τις C, C++.
- PyPy JIT (Just in Time): Επικεντρώνεται στην ταχύτητα, την αποτελεσματικότητα και τη συμβατότητα. Μετατρέπει το Python κώδικα σε γλώσσα μηχανής κατά την εκτέλεση.

- **Αποστολή:** Προώθηση, προστασία και πρόοδο της Python. Υποστήριξη και διευκόλυνση της ανάπτυξης μιας πλουραλιστικής διεθνούς κοινότητας προγραμματιστών της.
- Μπορεί όποιος επιθυμεί να συνεισφέρει σε έργο ή σε χρήμα
- Python Software Foundation License
- PyCon (επόμενο 20-26 Ιούλη στο Bilbao)
- Χρηματοδοτήσεις (python sprints)

Systems

25 Python jobs in Systems

NEW

Contract Python Developer -

Amsterdam, Netherlands

Broadcast

Frame 25 (Agency)

Looking for: Back end, Database, Integration, Systems, Web, python, celery, django, git, sqlalchemy

Posted: 16 April 2015

Category: Developer / Engineer

NEW

BioTech Python Software Engineer

Redwood City, CA, USA

Synthego

Looking for: Back end, BigData, Integration, Systems, Web, Robots, Lasers, Linux, Flask, Django, Pandas, Git, celery, ZMQ, CouchDB, Xen, Elasticsearch

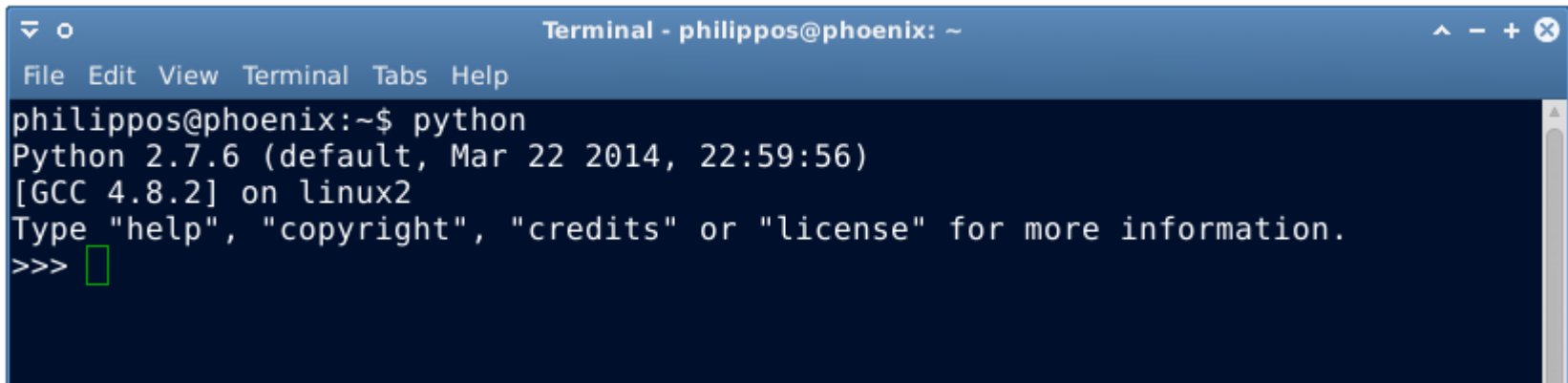
Posted: 16 April 2015

Category: Developer / Engineer

PYTHON

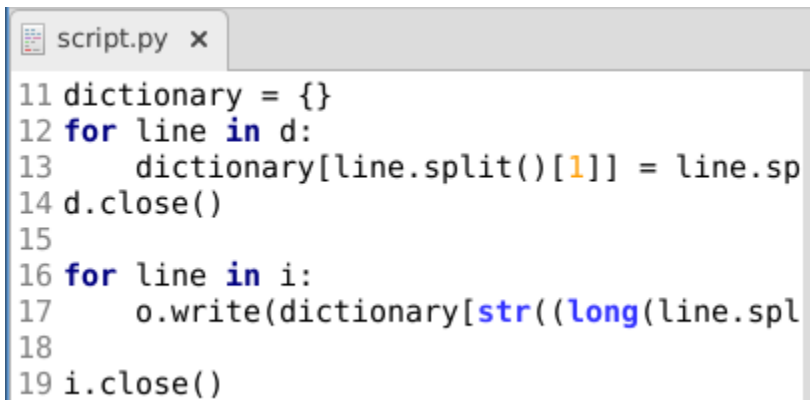
How to script?

- Interactively from terminal (interpreted language)

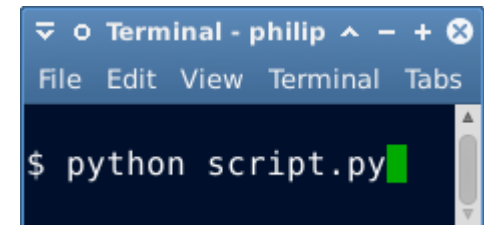
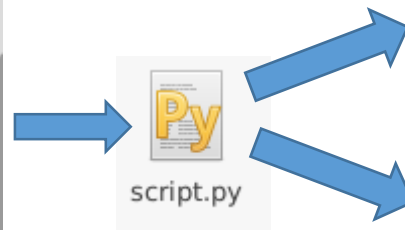


```
Terminal - philippos@phoenix: ~
File Edit View Terminal Tabs Help
philippos@phoenix:~$ python
Python 2.7.6 (default, Mar 22 2014, 22:59:56)
[GCC 4.8.2] on linux2
Type "help", "copyright", "credits" or "license" for more information.
>>> █
```

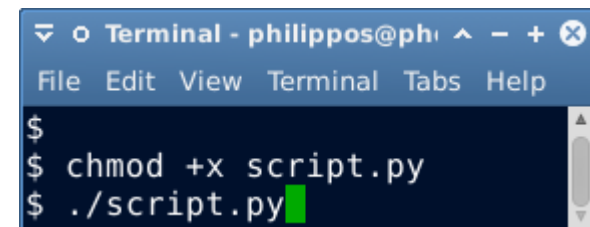
- From file



```
script.py x
11 dictionary = {}
12 for line in d:
13     dictionary[line.split()[1]] = line.sp
14 d.close()
15
16 for line in i:
17     o.write(dictionary[str((long(line.spl
18
19 i.close()
```



```
Terminal - philip  ^ - + X
File Edit View Terminal Tabs
$ python script.py █
```



```
Terminal - philippos@ph  ^ - + X
File Edit View Terminal Tabs Help
$
$ chmod +x script.py
$ ./script.py █
```


PYTHON

First Program in Python



Hello world

```
print "Hello World\n"
```

Simple operations

```
print 25+5, 25-5, 5*5, 25/5
```

Using variables

```
number = 2
```

```
other = 3
```

```
print 'Sum: {0}'.format(number + other)
```

PYTHON

Strings



```
mystring = 'Hello'  
mystring += ' world!'  
print mystring  
print len(mystring)
```

```
if <condition1>:  
    <statement1>  
elif <condition2>:  
    <statement2>  
else:  
    <statement3>
```

```
students = ['Luis', 'Mark', 'Rita']  
print students [0], students [1], students [2]
```

```
students = ['Luis', 'Mark', 'Rita',...]  
for st in students:  
    print st
```

```
values = [ 0.11, -0.23 ]  
print values [0], values [1]
```

PYTHON

List Indexing



```
students = ['Luis', 'Mark', 'Rita']
```

```
print students [0]
```

```
print students [0:1]
```

```
print students [1:]
```

```
print students [-1]
```

```
print students [-2]
```


PYTHON

Nested List



```
matrix = [  
    [1,2,3,4],  
    [5,6,378],  
    [9,10,11,12],  
]
```

- Basic types: int, float, list
- Control flow: for, while, if, else, elif

Τα tuples είναι σαν **αμετάβλητες** lists.

```
A = (0, 1, 2)
```

```
A = ( 1 )
```

```
print A[0]
```

```
print len(B)
```

Τα dictionaries είναι **associative arrays** όπου γίνεται indexing μέσω **keys**.

```
g = {}  
g['Nick'] = 18  
g['Rita'] = 20  
print g['Nick']  
print len(g)  
print g.keys()
```

Ένα set είναι ένα unordered collection χωρίς duplicate elements. Τα sets υποστηρίζουν και μαθηματικές πράξεις όπως η ένωση, η τομή και η διαφορά.

```
numbers = set( [1,2,5] )  
numbers.add(4)  
print numbers  
print numbers | set( ['Rita'] )  
print numbers & set( ['Rita'] )  
print numbers - set( [2,3] )
```



```
def greet1():  
    print 'Hello World'  
    print 'Still Here'  
def greet2(name):  
    print 'Hello {}'.format(name)
```

```
greet1()  
greet2('Luis')
```

```
class MyClass:  
    i = 10  
    def f(self):  
        return 'Hello World'
```

```
x = MyClass()  
print x.i  
print x.f()
```

```
try:
    10/0
except ZeroDivisionError:
    print 'Oops, invalid.'
else:
    print 'Pass.'
finally:
    print 'We \'re done with that.'
```

```
import random
```

```
from time import clock
```

```
num = random.randint(1,100)
```

```
print num
```

```
myfile = open('text.txt','w')  
myfile.write('This is a sample string')  
myfile.close()
```

```
myfile = open('text.txt','r')  
print myfile.read(20)  
myfile.close()
```


OverLap

Εισαγωγή

OverLap

OverLap

Βιβλιοθήκες, Παραδείγματα, Επίδειξη

- Ένα “Internet bot” που σερφάρει συστηματικά τον Παγκόσμιο Ιστό, συνήθως για σκοπούς Web Indexing.
- **Άλλα ονόματα:** Web Spider, Ant, Automatic Indexer ή Web Scutter.
- Οι μηχανές αναζήτησης χρησιμοποιούν Web Crawling για ενημέρωση των περιεχομένων ιστού ή των ευρετηρίων ιστοσελίδων.

- Μπορούν να αντιγράψουν σελίδες που επισκέφτηκαν για μεταγενέστερη επεξεργασία από Μηχανές Αναζήτησης.
- Έτσι η αναζήτηση γίνεται πιο αποδοτική για τους χρήστες.
- Μπορούν να επικυρώσουν hyperlinks και HTML κώδικα.
- Χρησιμοποιούνται επίσης και για web scraping (τεχνική εξαγωγής πληροφοριών από ιστοσελίδες).

- **Beautiful Soup** είναι ένα πακέτο της Python για parsing HTML και XML αρχεία.
- Δημιουργεί ένα parse tree για parsed σελίδες.
- Μπορεί να χρησιμοποιηθεί για εξώρυξη δεδομένων από HTML, κάτι πολύ χρήσιμο για Web Scraping.
- Διαθέσιμο για Python 2.6+ και Python 3.

OverLap

Beautiful Soup - Παράδειγμα



```
BS4Sample.py x
# anchor extraction from html document
from bs4 import BeautifulSoup
import urllib2

webpage = urllib2.urlopen('http://en.wikipedia.org/wiki/Main_Page')
soup = BeautifulSoup(webpage)
for anchor in soup.find_all('a'):
    print(anchor.get('href', '/'))
```

Run BS4Sample

```
/
#mw-head
#p-search
/wiki/Wikipedia
/wiki/Free_content
/wiki/Encyclopedia
/wiki/Wikipedia:Introduction
/wiki/Special:Statistics
/wiki/English_language
/wiki/Portal:Arts
/wiki/Portal:Biography
/wiki/Portal:Geography
/wiki/Portal:History
/wiki/Portal:Mathematics
/wiki/Portal:Science
/wiki/Portal:Society
/wiki/Portal:Technology
/wiki/Portal:Contents/Portals
```

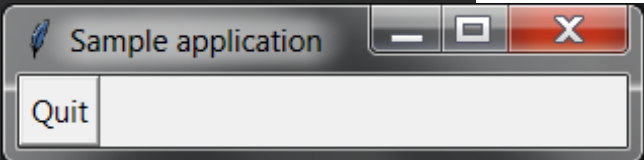
- Python's de-facto standard GUI (Graphical User Interface) package.
- Tkinter γράφτηκε από τον Fredrik Lundh.

```
sample.py x
#!/usr/bin/env python3
import tkinter as tk

class Application(tk.Frame):
    def __init__(self, master=None):
        tk.Frame.__init__(self, master)
        self.grid()
        self.createWidgets()

    def createWidgets(self):
        self.quitButton = tk.Button(self, text='Quit', command=self.quit)
        self.quitButton.grid()

app = Application()
app.master.title('Sample application')
app.mainloop()
```

A screenshot of a Tkinter application window titled "Sample application". The window has a standard Mac OS-style title bar with minimize, maximize, and close buttons. Inside the window, there is a single button labeled "Quit". The background of the window is dark, matching the code editor behind it.

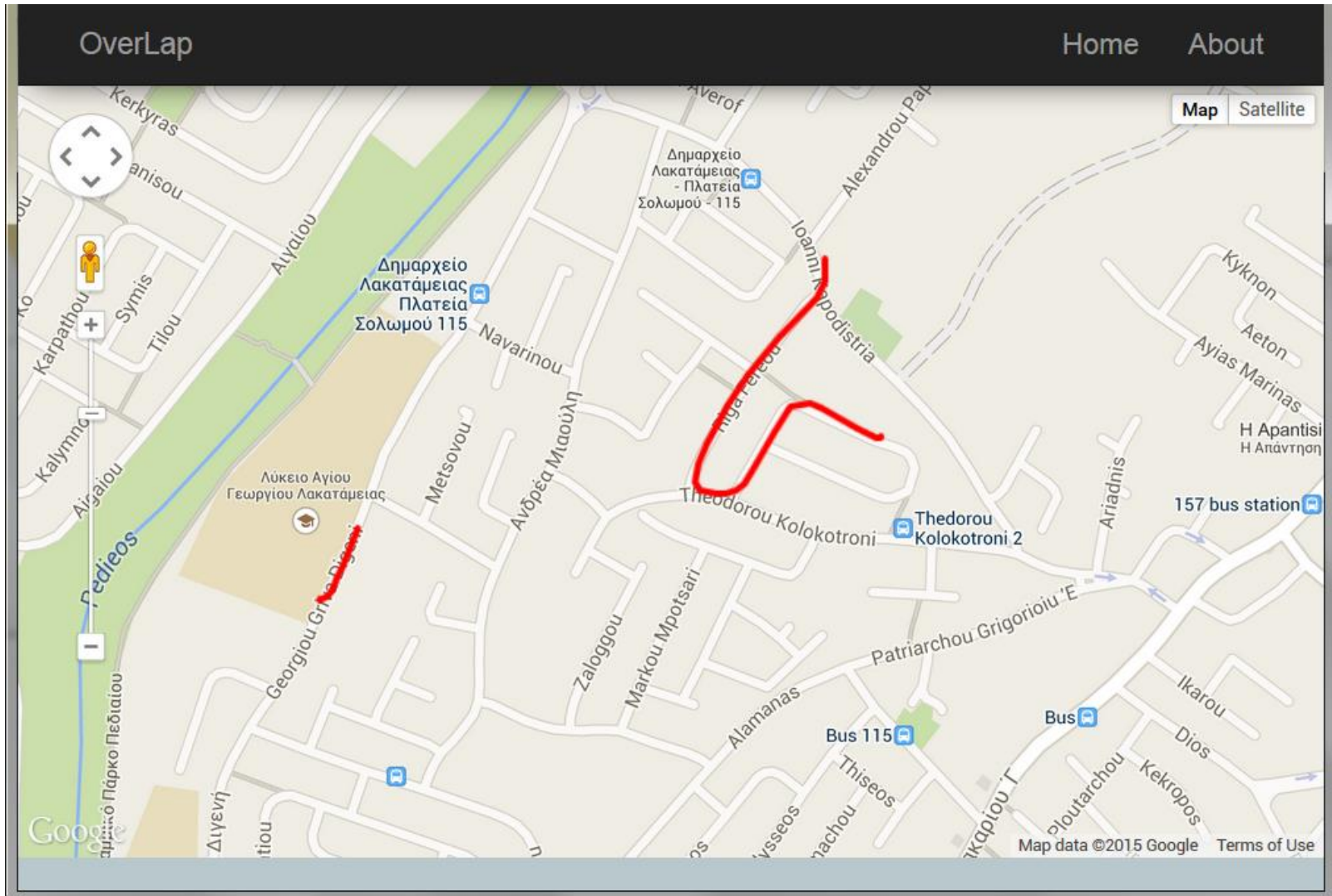
OverLap



```
> Crawling Runtastic...
> Logging in...Done!
> Accessing routes...Done!
> Downloading 440345311.gpx...Done!
> Saving Runtastic/andreas-andreou-4-/440345311.gpx...Done!
> Saving Default/440345311.gpx...Done!
> Downloading 440345314.gpx...Done!
> Saving Runtastic/andreas-andreou-4-/440345314.gpx...Done!
> Saving Default/440345314.gpx...Done!
> Crawling Runtastic...Done!
```


OverLap

Επίδειξη (backup slide)



Tracing Roads

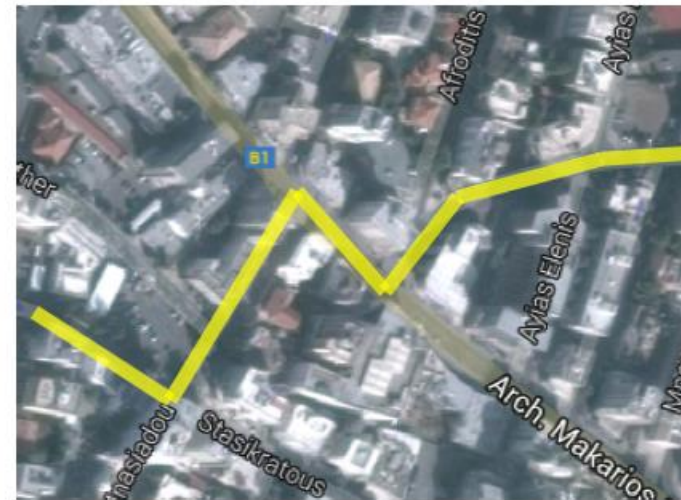
Algorithm

OverLap

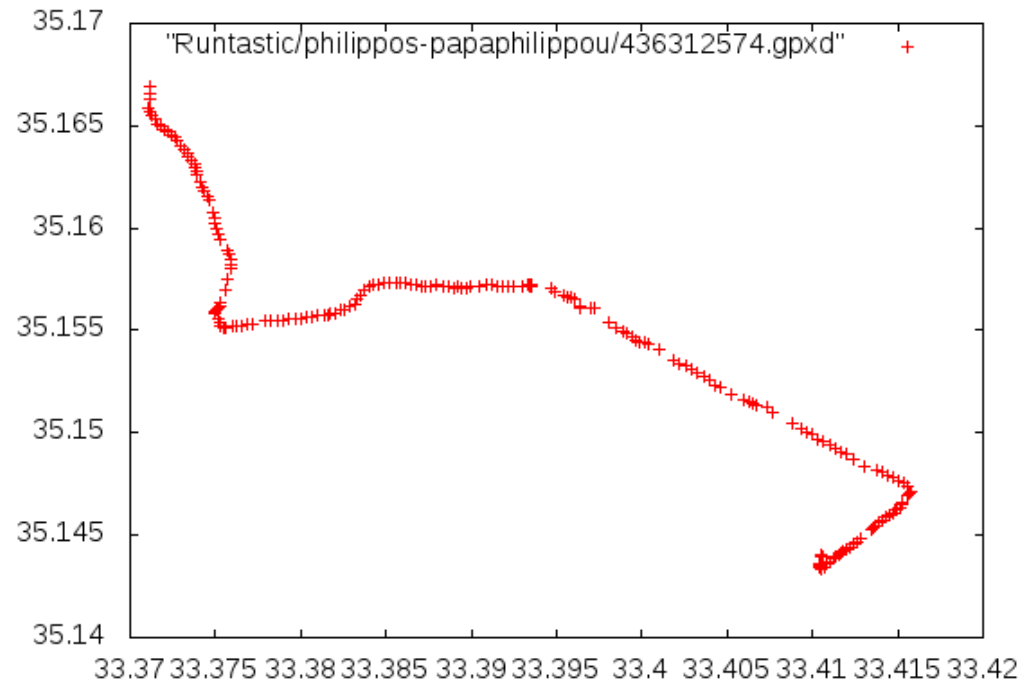
Tracing Roads (1)



- When combining many GPS traces of many users huge amount of data will build up.
- We need an algorithm that distinguishes roads to keep only 2 points for each road pass.



- Similar techniques are present in image compression libraries
- Real world problem example
 - Small Distance
 - Hundreds of points
 - Only for 1 trip
- The problem scales
- More problems when previewing in browser

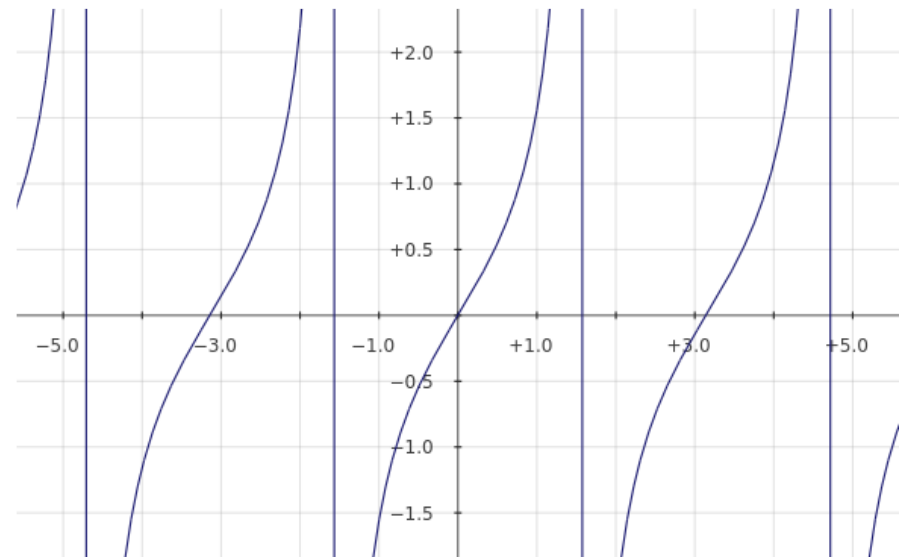
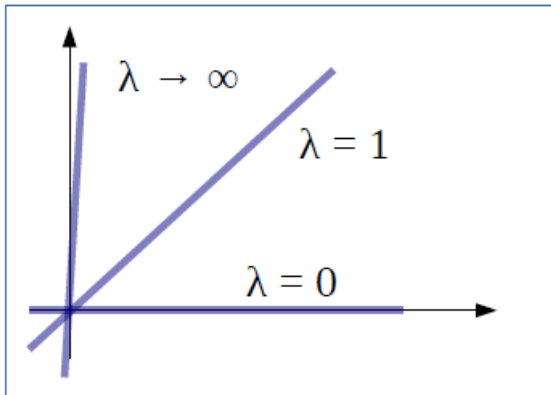


- Algorithm (pseudo-code):
 - For each collected GPX file (trip)
 - For each point traveled
 - Compute gradient between the current and previous point
 - If much different
 - Save previous road
 - Else
 - Continue computing the current road

- Finding the gradient

- We initially use the formula:
$$\lambda = \frac{\psi_1 - \psi_2}{\chi_1 - \chi_2}$$

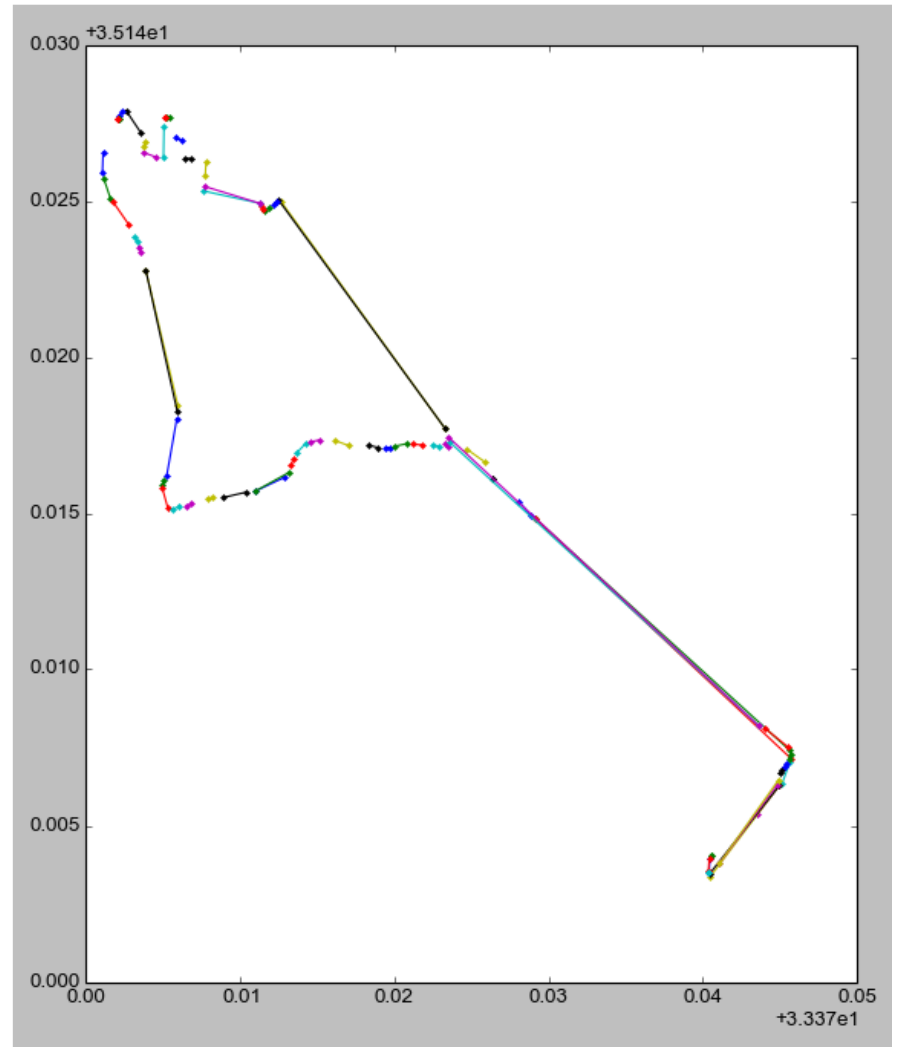
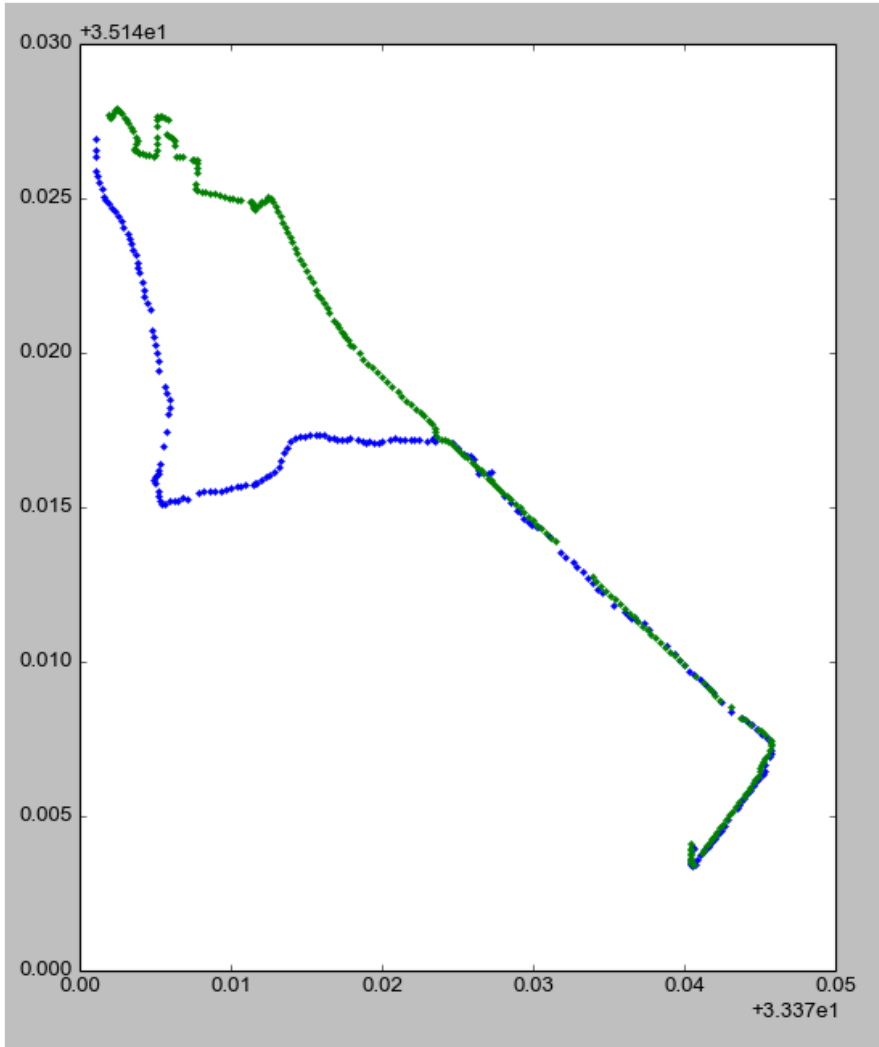
- Gradient difference ($\lambda_2 - \lambda_1$)
 - Gradient difference alone isn't sufficient for our case



- We need difference of bearings
$$d = | \arctan(\lambda_2) - \arctan(\lambda_1) |$$

OverLap

Tracing Roads (5)

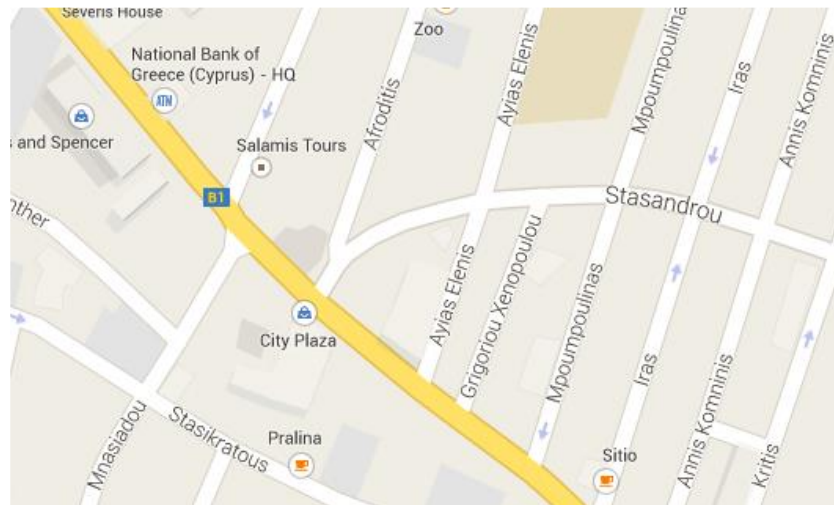
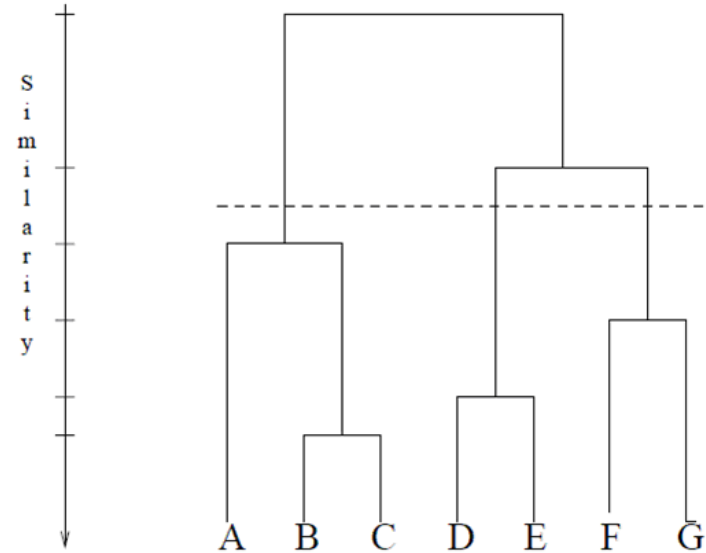


OverLap

Tracing Roads (6)



- Alternative methods
 - Hierarchical clustering (topic of Data Mining)
- Roads specified by service e.g. Google Maps



Πηγές

Πηγές, tutorials...



- <http://www.stavros.io/tutorials/python/>
- <https://www.python.org/>
- https://en.wikipedia.org/wiki/Python_%28programming_language%29
- <http://pypy.org/>
- http://page.mi.fu-berlin.de/prechelt/Biblio/jccpprt_computer2000.pdf
- <http://www.explainkcd.com>

- **Year 10 Interactive Maths - Second Edition**

http://www.mathsteacher.com.au/year10/ch03_linear_graphs/02_gradient/line.htm

- **Data Mining Algorithms In R/Clustering/Hybrid Hierarchical Clustering**

http://en.wikibooks.org/wiki/Data_Mining_Algorithms_In_R/Clustering/Hybrid_Hierarchical_Clustering

Ερωτήσεις?

Ευχαριστούμε για την προσοχή σας