ASSIGNMENT GUIDELINES

INFO-F-438

Course assistant: Sofia Papadimitriou

Overview

- General type of assignments
- Style of files and code
- Grading and requirements
- Plagiarism

Overview

- General type of assignments
- Style of files and code
- Grading and requirements
- Plagiarism

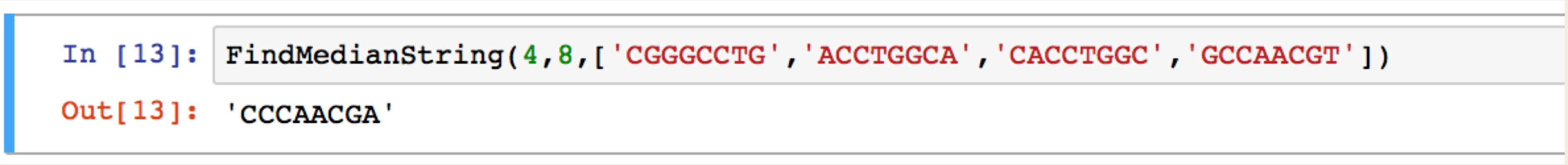
You are requested to write a program…

>> Write a **program** that solves the **median string problem** on the alphabet $\{A, C, G, T\}$.

… with a specific input and output…

>> This program should take three parameters"
 t: number of sequences,
 l: length of sequences,
 and a collection **C** of t strings in $\{A,C,G,T\}^{l}$

and **return the median string c** in $\{A,C,G,T\}^{l}$.



```
In [13]: FindMedianString(4,8,['CGGGCCTG','ACCTGGCA','CACCTGGC','GCCAACGT'])
Out[13]: 'CCCAACGA'
```

... and specific subproblems to consider

>> For this purpose, you should think about:

1. a way to **search the solution space** $\{A, C, G, T\}^l$

2. a simple **branch-and-bound** technique to quickly prune uninteresting branches

3. a **simple way to initialize the search** so that pruning is likely to be efficient.

Overview

- General type of assignments
- Style of files and code
- Grading and requirements
- Plagiarism

Types of program files

ipython Notebook

PREFERRED

You can make the assignment as a report.

You can try different algorithms in the same file (provided that you explain what you do).

Types of program files

python script

OKAY

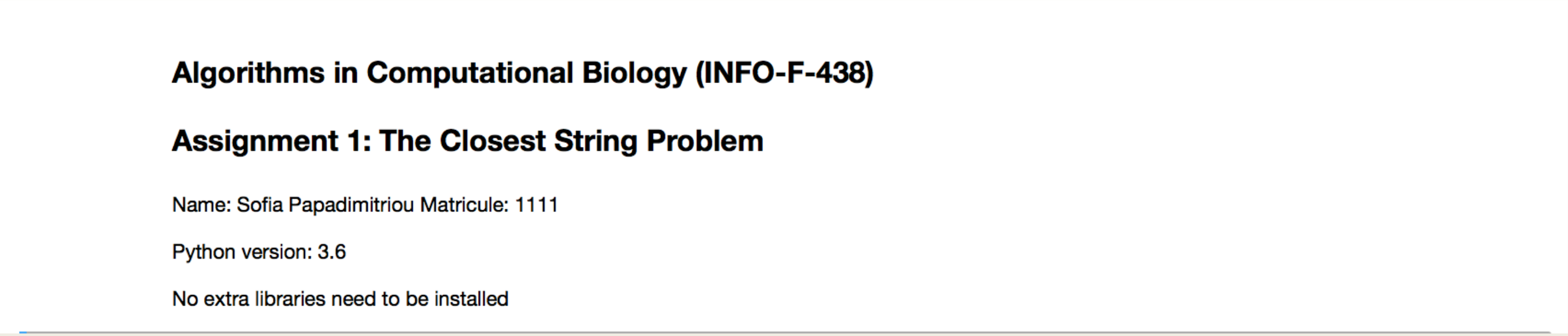
Provide enough **instructions on how to run it.**

Well documented.

Different algorithms? Different files.

Starting your code

Give me enough documentation before testing your code.



Algorithms in Computational Biology (INFO-F-438)

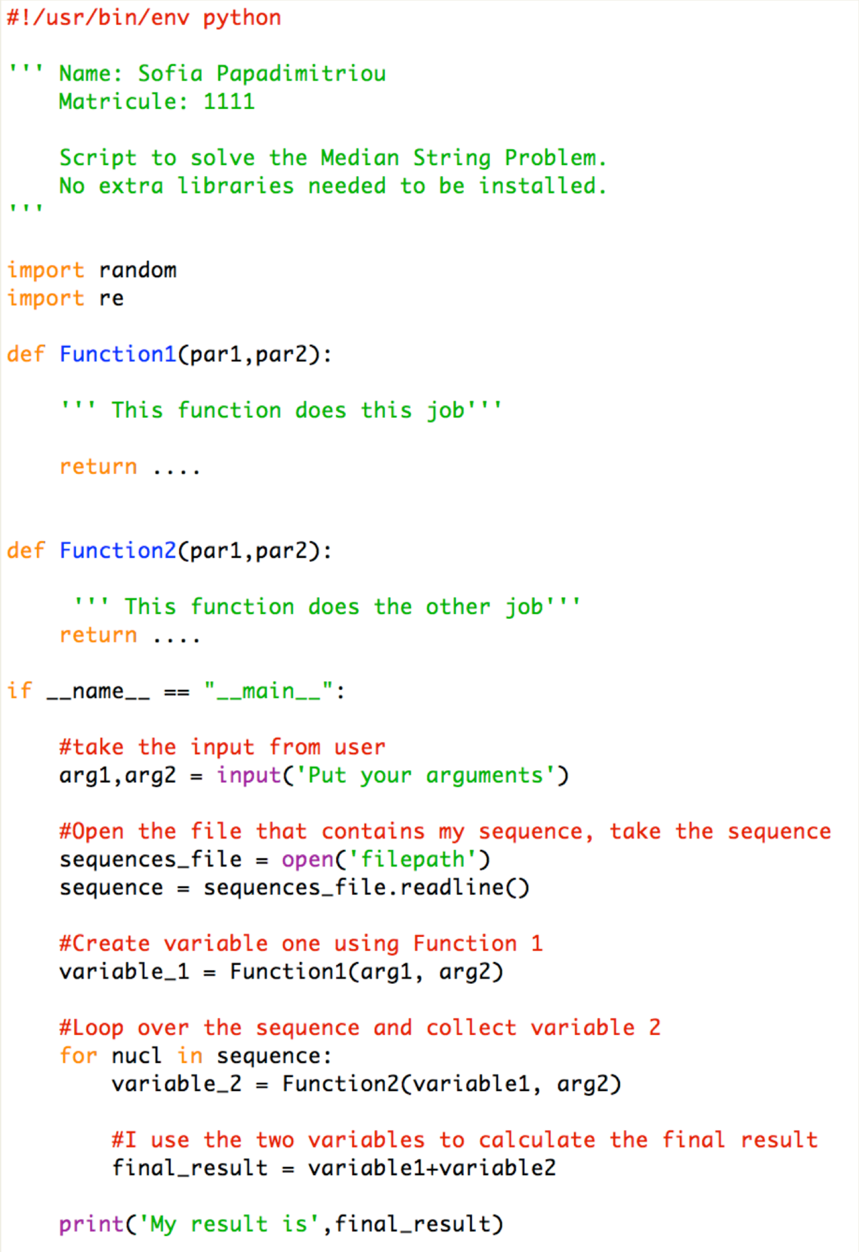
Assignment 1: The Closest String Problem

Name: Sofia Papadimitriou Matricule: 1111

Python version: 3.6

No extra libraries need to be installed

Organize your code!



```python
#!/usr/bin/env python

''' Name: Sofia Papadimitriou
    Matricule: 1111

    Script to solve the Median String Problem.
    No extra libraries needed to be installed.
'''

import random
import re

def Function1(par1,par2):

    ''' This function does this job'''

    return ....


def Function2(par1,par2):

     ''' This function does the other job'''
    return ....

if __name__ == "__main__":

    #take the input from user
    arg1,arg2 = input('Put your arguments')

    #Open the file that contains my sequence, take the sequence
    sequences_file = open('filepath')
    sequence = sequences_file.readline()

    #Create variable one using Function 1
    variable_1 = Function1(arg1, arg2)

    #Loop over the sequence and collect variable 2
    for nucl in sequence:
        variable_2 = Function2(variable1, arg2)

        #I use the two variables to calculate the final result
        final_result = variable1+variable2

    print('My result is',final_result)
```

Initial Documentation

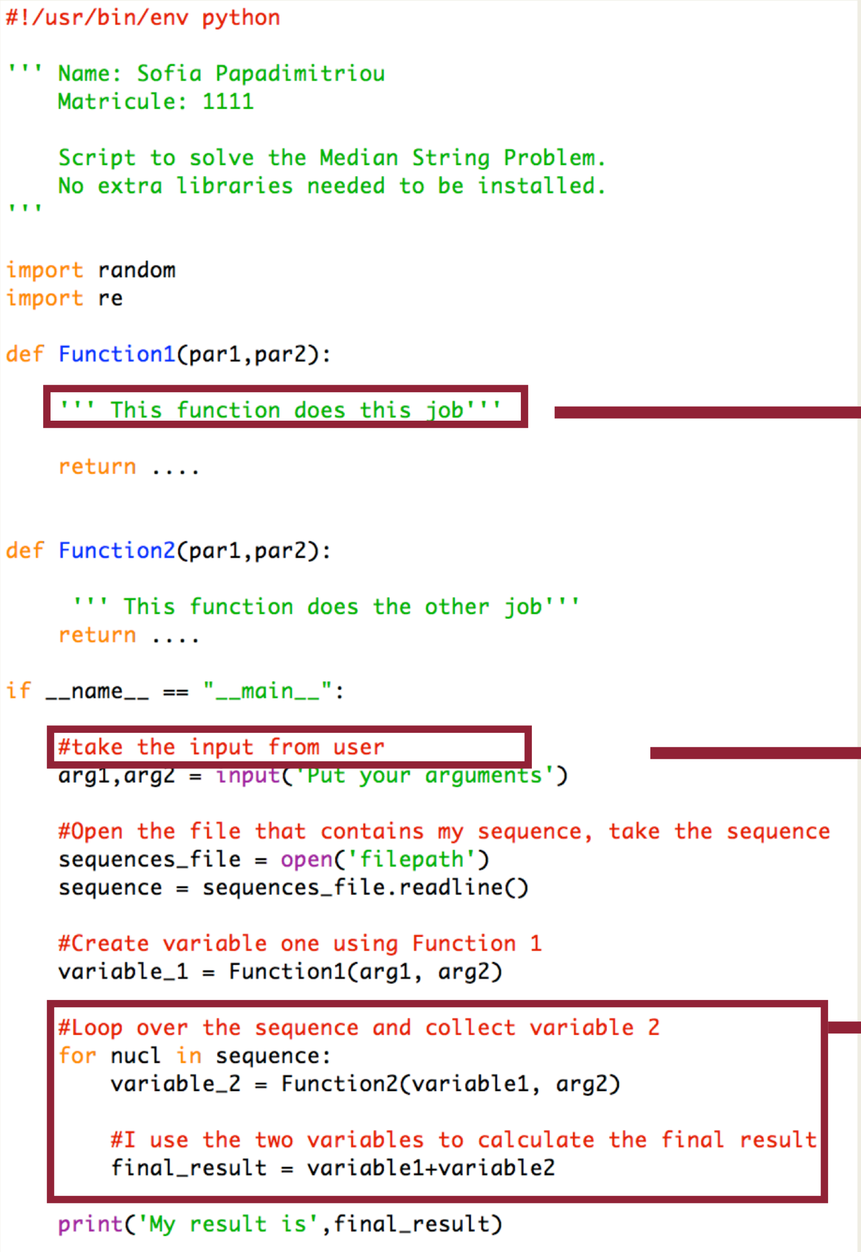
Imports

Functions

Main code:

user input, files, use of functions, further development

Document your code!



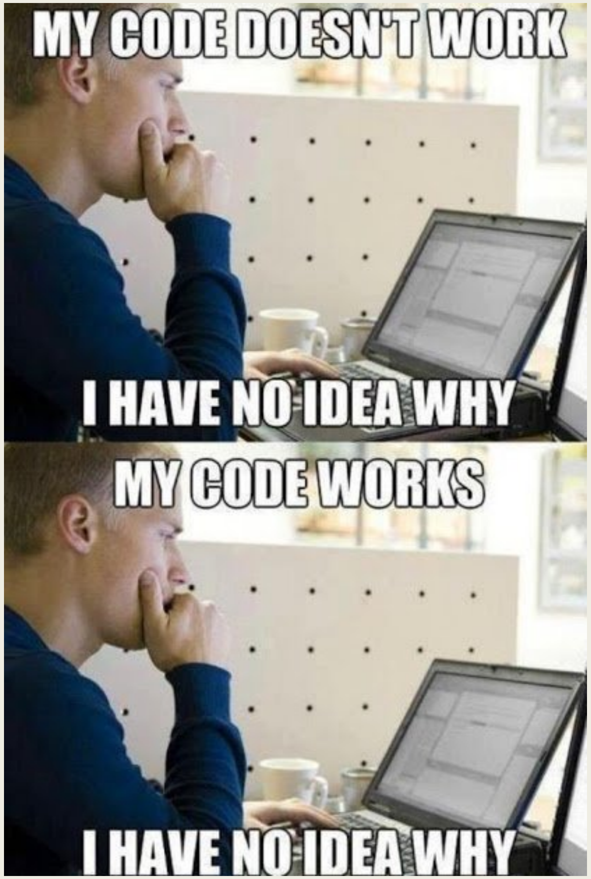
```
#!/usr/bin/env python

''' Name: Sofia Papadimitriou
    Matricule: 1111

    Script to solve the Median String Problem.
    No extra libraries needed to be installed.
'''

import random
import re

def Function1(par1,par2):

    ''' This function does this job'''

    return ....


def Function2(par1,par2):

     ''' This function does the other job'''
    return ....

if __name__ == "__main__":

    #take the input from user
    arg1,arg2 = input('Put your arguments')

    #Open the file that contains my sequence, take the sequence
    sequences_file = open('filepath')
    sequence = sequences_file.readline()

    #Create variable one using Function 1
    variable_1 = Function1(arg1, arg2)

    #Loop over the sequence and collect variable 2
    for nucl in sequence:
        variable_2 = Function2(variable1, arg2)

        #I use the two variables to calculate the final result
        final_result = variable1+variable2

    print('My result is',final_result)
```

What does the function do?

Justify your actions

Justify a block of code

Test your code!



Does the code take user input correctly?

Test your functions with examples:
does the function give the expected output?

Check for syntax errors:
forgotten parenthesis, inserts

RUN the complete file before sending your work to check for errors.

Overview

- General type of assignments
- Style of files and code
- Grading and requirements
- Plagiarism

Who takes the 10?

To take the **maximum grade (10)**:

- you should implement the correct algorithm that we ask for
  e.g. correct implementation of branch-and-bound

- you should implement the improvements that are asked in your assignment
  e.g. a more efficient starting point of the algorithm

- your code should work for a selection of different cases.
  e.g. different number of sequences, small/long sequences, AAAAAAAA vs CCCCCCC, sequences of different length

Grading criteria

○ Correctness of code and output

○ Efficiency of the algorithm

○ Generality of algorithm

○ Documentation

○ Respect of instructions

○ Errors in code

○ Submission on time!

Overview

- General type of assignments
- Style of files and code
- Grading and requirements
- Plagiarism

Plagiarism in terms of coding

You take a code from another source (complete or parts of it) and you provide it in your assignment as your own **without explicitly specifying that source**

Plagiarism in terms of coding

You take a code from another source (complete or parts of it) and you slightly modify it:
e.g. variable names, order of lines, comments

and you provide it in your assignment as it is **without explicitly specifying that source**

Plagiarism in terms of coding

You provide someone else’s code as your own, even if that person has agreed.