

# Introduction to Python - Python II

## 1 Getting Started

Before you start with the following assignments, please create a directory `PythonLab2` in your home directory (`~/`). In this directory save all the files you create during this assignment.

## 2 Assignmet - Similarity of Sequences

Write in an editor the program, which calculates the distance between two sequences `A = "ACGT"` and `B = "AGGT"`.

A simple program (without function and modules) is sufficient.

1. Calculate the distance between the following sequences and print out the result. Since the following sequences are already aligned, we can calculate the distance between them. Change your program so that it can read two aligned sequences from the command line. Test your program with the following sequences.
  - a) `ACGT` and `A-GT`
  - b) `AC-GT` and `AGT--`
  - c) `AC-CGT` and `AGT---`
  - d) `ACCGT` and `TGCCA`
  - e) `GATT-ACA` and `TACCATAC`
  - f) `--GA--TT--AC-A` and `TA--CC--AT--CA`
2. Extend the program that the aligned sequences are printed out additionally to their distance.
3. Extend the program that the distance between two sequences is only calculated when both sequences have the same length. Test your program with the input sequences:
  - a) `ACGT` and `AGT`
  - b) `ACCGT` and `TGCCA`

4. Extend the program that the second sequence is inverted and assigned to a third sequence. Please, read the first and second sequence from the command line. Calculate the distances between the first and the second and between the first and the third sequence. Compare the distance between the first and the second and the first and the third sequence and print the alignment with the smaller distance. If the distances are equal, then print the alignment of the first and second sequence.

Test your program with the following sequences:

- a) ACGT and A-GT
- b) AC-GT and AGT--
- c) ACCGT and TGCCA
- d) GATT-ACA and TACCATAC