Entry level computer-programming for bio-medical students

The “big data” phenomena entered the biomedical research as well as into the medical practice. Microarray or NexGen sequencing experiments can produce mass data well beyond the capabilities of direct human analysis. Personalized medicine approach requires the intensive use of biomedical databases i.e. accessing the “big data”. These challenges demand new data handling and analysis competences from the researchers and biomedical students. Data handling can be done with the aid of dedicated tools (commercial or GPL) but these solutions are typically limited to the data from a certain instrument of a certain vendor. Computer programming could provide a way for restriction free analysis of the data from different sources. Programming approach to data handling makes the research more effective and creative.

The aim of the course is the presentation of basic programming skills to the students as a “starter kit” witch they can develop further on their own by the aid of the online materials according to the needs of their scientific projects and interests. The course starts from the novice level and no previous programming experience needed.

The course is concentrating on the Python programming language. Python is common in bioinformatics and in fact across natural sciences, under GPL and ideal for novice programmers due to its clear and simple structure but it has the potential of powerful solutions at a later stage of learning.

The course is problem oriented. The students learn Python via examples starting with simple ones and analyzing more complex, real-life like biological data at later stage. Every class ends with a few exercises to solve by the next occasion and the solution is discussed at the beginning of the next class. The students are encouraged to solve the problems in groups and real-life data from their daily lab work is also welcomed. The final evaluation of their progress is based on the solutions of exercises they are handing in during the course.

The course is 14 classes, one per week, 90 minutes each (2 credits). The students are asked to bring their own computer to use. The classes held by Dr Miklos Cserzo, Dr Bence Szalai and Dr Gabor Turu with the following titles:

1. The Anaconda and the Jupyter systems, variables and operators
2. File access, I/O handling
3. Lists and loops
4. Branching and conditional statements
5. Dictionaries and functions
6. Summing up the basics
7. Practicing the basics
8. Libraries – matplotlib I
9. Libraries – pandas II
10. More on libraries
11. Practicing I
12. Practicing II
13. Practicing III
14. Python extras – list comprehension, classes