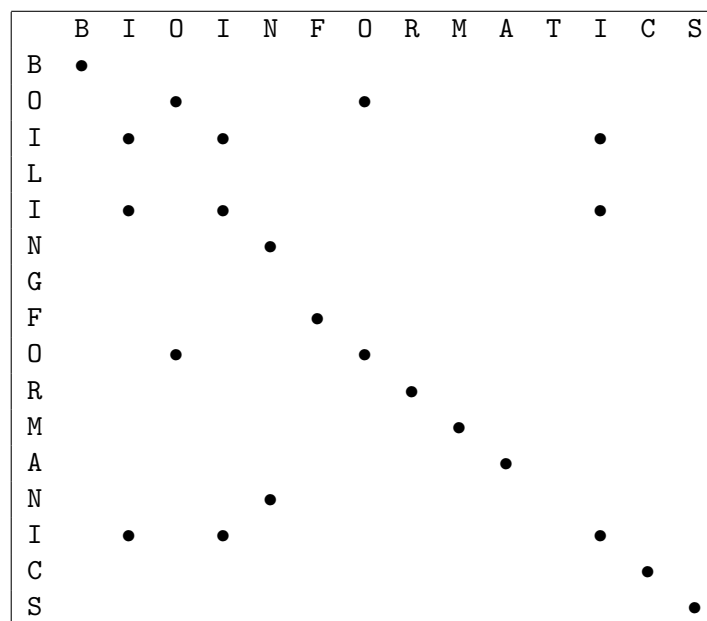


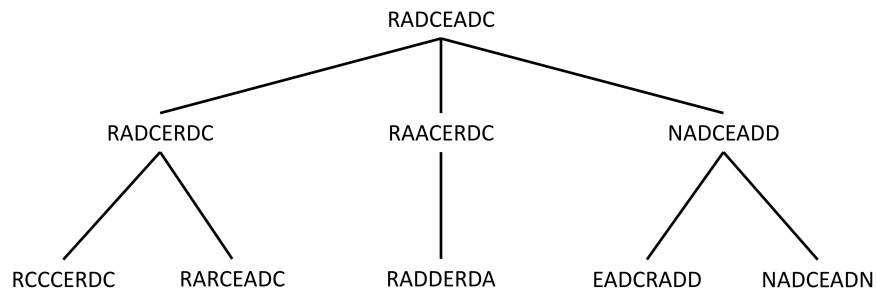
Assignment no. 4

Exercise 8 (12 points) A dot plot is a graphical method that can be used in order to compare two biological sequences. It allows to identify regions of close similarity between them. See example below. It shows a dot plot for the (non biological) strings BIOINFORMATICS and BOILINGFORMANICS.



Write a Python program that takes two strings (not necessarily of the same length) from the command line and plots the dot plot for these strings on the standard output. Document and test your program in detail and provide checks for the command line arguments.

Exercise 9 (12 points) Given the (imaginary) phylogenetic tree for nine amino acid sequences over the alphabet $\{A, C, D, E, N, R\}$ compute (on paper) the PAM1 matrix as described in the lecture notes. Derive also the corresponding scoring matrix. How would you obtain a PAM16



matrix from this PAM1 matrix?

Submission: electronically via Moodle:

<https://moodle.jku.at>

Deadline: November 28, 2019, 14:00.