Project - Notebook

DD2404 - Bioinformatics
Reducing noise in protein multialignments

Christopher Herron - cherron@kth.se - 1993.09.20-1979
David Olanders - davola@kth.se - 1993.06.04-8613
André Zachrisson - andrez@kth.se - 1994.05.21-8257

Examiner: Lars Arvestad

Project Notebook

For this project a Notebook has been kept for every day that the project was worked on.

17 December

We initiated the project with reading all the assigned material and setting up the structure for what software's we wanted to use. We decided to create a folder in GitHub with a version control software that we could link to Atom. We also decided to use Overleaf as an LaTeX online editor for the report as well as for the notebook of the project. After this we started writing the script for the noise reduction. The script is fully functioning but requires some more testing to be sure that the output is correct:

What needs to be done:

- Print output tree
- Work with Dendropy

18 December

We expanded the script so that we now can run all input files via a temporary directory. Also, we can now save all distances for both the reduced and original alignments, though with a slow script.

During the work described above, we had contact via email with the examiner Lars Arvestad. We asked about what to use FastPhylo and DendroPy for since there arose some confusion about the assigned software. After this contact the confusion was cleared and we could expand the script as described above.

What needs to be done:

- Calculate averages for the different types of alignments
- Plan presentation of collected data
- Finalize and optimize script

19 December

We finalized the script and worked on optimizing it, however since we use subprocess it is still very slow. Furthermore we had some time consuming issues with the final parts of the script, especially with getting the statistics. The report has been structured and the writing has started.

What needs to be done:

- Finishing the report
- Structure and comment code for others use

20 December

Finished the project.