

Second Semester Report

Lénárd Lajos Szánthó (lenard.szantho@ttk.elte.hu)

Doctoral School of Physics

Statistical Physics, Biological Physics
and Physics of Quantum Systems Program

Supervisor: Gergely J. Szöllősi

Thesis title: Developing next-generation phylogenetic methods

June 22, 2022

Introduction

Phylogenetics studies the evolutionary relationships between organisms by reconstructing their tree of descendants, called a phylogeny. During my doctoral studies, I am developing new methods and applications as described in the first semester's report.

Description of research work carried out in current semester

In the beginning of this semester the manuscript describing our results in the project *New method to detect and ameliorate long branch attraction (LBA) artefacts: CAT-PMSF* was finished and submitted to the journal Systematic Biology. The review process has begun and early June we got positive feedback: *accepted with minor revision*. Currently I am revising the article and hope to submit it by middle of July.

The implementation of *Horizontal gene transfer highways* was continued, currently there are two tools written in Python, one simulating the evolution in forward direction and another (the original one) tracing it back. This allows us to test the predictions against simulations. Further work was devoted to implement the Conditional Clade Probabilities, which allows the software to consider many tree topologies with their appropriate probability weights during estimating the likelihood of the gene trees given the species tree.

The *Date the tree of fungi* project preliminary trees were obtained which had bogus topology (some species of the taxon Zoopagomycota were clustering with another taxon Mucoromycota), we were checking whether the sequence data may have been contaminated, but it is not, more sophisticated model needs to be used to describe the data.

Publications

The manuscript for the project *New method to detect and ameliorate long branch attraction (LBA) artefacts: CAT-PMSF* was accepted with minor revision by the journal Systematic Biology, the revision is ongoing as of writing this report.

Studies in current semester

I have enrolled to the following classes:

BIO/02/01E Behaviour genetics (Enikő Kubinyi)

BIO/10/32G Computer Modelling in Biology (Viktor Müller)

FIZ/3/089 Deep learning and machine learning in natural sciences (Biricz András Mátyás, Csabai István Ervin, Olar Alex, Pataki Bálint Ármin, Udvarnoki Zoltán András)

The Behaviour genetics and Computer Modelling in Biology classes were graded with 5. Sadly, due to personal reasons, I did not have time to actively participate on the Deep learning and machine learning in natural sciences class, I plan to retake the course next spring.

Conferences in current semester

I have not attended any conferences this semester.

Teaching activities

I have taught 3 computational practical classes for the course *Statistical Physics B*, *elmfiz4bf19va* for BSc students.

Professional activities

I have helped in the development of the Kooplex Research and Teaching System, which provides on-demand JupyterNotebook access to the students and researchers of the university (or collaborators of), although, due to personal reasons I was not very active after April.

I have maintained the research group's high performance computing (HPC) cluster.

I have helped to maintain the university's high performance computing (HPC) cluster which service is provided by the ELTE IIG.