

Exam DM813 Algorithms for Biological Sequence Analysis (2010)

Time and Place

Time: Thursday, January 20, 2010.

Place: The exam takes place in Auditorium U49.

The starting time for the first examination is still to be determined. A mail with this information will be sent as soon as the external examiner agreed.

Even though the expected total examination time per student is 30 minutes (see below), it is not possible to calculate the exact examination time from the placement on the list, since students earlier on the list may not show up. Thus, students are expected to show up plenty early. In principle, all students who are taking the exam on a particular date are supposed to show up when the examination starts, i.e., at the time the first student is scheduled. This is partly because of the way external examiners are paid, which is by the number of students who show up for examination. For this particular exam, we do not expect many no-shows, so showing up one hour before the estimated time of the exam should be safe.

Procedure

The exam is in English. When it is your turn for examination, you should use approx. 10 minutes to defend your project. You can either cover Mandatory Assignment 1: Dynamic Programming and Sequence Alignment, or Mandatory Assignment 2: Cophylogenetics (not both). Note that you have no preparation time. We, the examiner and the censor, will supplement with specific questions when appropriate, and after a while, we will end the discussion of the defence. After that we turn to material from other parts of the curriculum. This part lasts approximately 20-22 minutes. The list of questions can be found below. Aim for a reasonable high pace and focus on the most interesting material related to the question. You are not supposed to use note material, textbooks, transparencies, computer, etc.

Note that all of this as well as discussion between examiner and censor about the grade will take about 30 minutes, so do not count on more than 9-11 minutes for your own presentation.

Some of the questions below are very broad, so you must select the material you choose to cover. You will of course also be evaluated based on your selection of material. If you only present the simplest material, you limit the grade you can obtain. On the other hand, a good presentation of the simple material is better than a very poor presentation of the harder material. We might of course still ask you questions about material that you have decided to skip.

Curriculum

Note that the curriculum given below is a subset of the curriculum given in the official course description.

The curriculum consists of

- Chapters 1, 4.1-4.5, 5.1-5.4, 6.1-6.2, 6.4, 7.1-7.2, and 8.1-8.4 from the book "Understanding Bioinformatics", 2008, by Marketa Zvelebil and Jeremy O. Baum. ISBN-13: 978-0815340249
- Chapter 1 from "Tangled Trees - Phylogeny, Cospeciation, and Coevolution", 2003, edited by Roderic D. M. Page. ISBN-13: 978-0226644677
- Pages 1-23 from Combinatorics of Genome Rearrangements, 2009, by Guillaume Fertin, Anthony Labarre, Irena Rusu, Eric Tannier and Stéphane Vialette. ISBN-13: 978-0262062824
- A parameter-adaptive dynamic programming approach for inferring cophylogenies. Daniel Merkle, Nicolas Wieseke, Martin Middendorf, BMC Bioinformatics 2010 (11).
- Solving the Preserving Reversal Median Problem, M. Bernt, D. Merkle, M. Middendorf, IEEE/ACM Transactions on Computational Biology and Bioinformatics, Vol. 5, No. 3, 332-347, 2008.
- A faster fixed-parameter approach to drawing binary tanglegrams. Sebastian Boecker, Falk Hueffner, Anke Truss and Magnus Wahlstrøm. (ALGO 2009).

Questions

1. Dynamic programming for pairwise alignment of DNA and protein sequences
2. BLAST and FASTA
3. BLOSUM and PAM matrices
4. Suffix tree and its applications
5. Methods for constructing evolutionary trees, e.g., Fitch Margoliash, Neighbor Joining, UPGMA, Parsimony
6. HMMs for multiple sequence alignment
7. Maximum Likelihood Method
8. Dynamic programming for inferring coevolutionary history of groups of species
9. Genome rearrangements based on PQ-trees