582606 Introduction to bioinformatics

Autumn 2007
Esa Pitkänen
Master's Degree Programme in Bioinformatics (MBI)
Department of Computer Science, University of Helsinki
http://www.cs.helsinki.fi/mbi/courses/07-08/itb/
Administrative issues

- Master level course
- Obligatory course in the Master’s Degree Programme in Bioinformatics
- 4 credits
- Prerequisites: basic mathematical skills
- Lectures: Tuesdays and Fridays 14-16 in Exactum C222
- Exercises: Wednesday 14-16 in Exactum C221
Teachers

- Esa Pitkänen, Department of Computer Science, University of Helsinki
- Prof. Elja Arjas, Department of Mathematics and Statistics, University of Helsinki
- Prof. Samuel Kaski, Laboratory of Computer and Information Science, Helsinki University of Technology
- Lauri Eronen, Department of Computer Science, University of Helsinki
How to enrol for the course?

1. Use the registration system of the Computer Science department: [https://ilmo.cs.helsinki.fi](https://ilmo.cs.helsinki.fi)

2. If you don’t have a student number or Finnish id yet, don’t worry: attend the lectures and exercises, and register when you have the id
How to successfully pass the course?

- You can get a maximum of 60 points
  - Course exam: maximum of 50 points
  - Exercises: maximum of 10 points
    - 0% completed assignments gives you 0 points, 80% gives 10 points, the rest by linear interpolation
    - “A completed assignment” means that you are willing to present your solution to the class in the exercise session
- Course will be graded on the scale 0-5
  - To get the lowest passing grade 1/5, you need to have at least 30 points
- Course exam: Wednesday 17.10. at 16.00-19.00 in A111
Course material


- Available at Kumpula and Viikki science libraries; book stores (Amazon.com ~$56, Akateeminen kirjakauppa ~75€, Yliopistokirjakauppa 71€)

- It is recommended that you have access to the course book!

- Slides for some lectures will be available on the course web page (copies in room C127)
Additional material

- Check the course web site
- N. C. Jones & P. A. Pevzner: An introduction to bioinformatics algorithms
- Alberts et al.: Molecular biology of the cell
- Lodish et al.: Molecular cell biology
Course contents

- Biological background (book chapter 1)
- Probability calculus (chapters 2 and 3)
- Sequence alignment (chapter 6)
- Rapid alignment methods: FASTA and BLAST (chapter 7)
- Phylogenetic trees (chapter 12)
- Expression data analysis (chapter 11)
Master's Degree Programme in Bioinformatics (MBI)

- Two-year MSc programme
  - You need to have your Bachelor’s degree ready by August 2008
MBI programme

• MBI educates bioinformatics professionals who
  – Specialise in computational and statistical methods
  – Work in R&D tasks in biology and medicine
MBI programme

• Two-year masters programme (120 cr)
• Offered jointly by the University of Helsinki (HY) and Helsinki University of Technology (TKK)
• Began in 2006 as a national programme, 2007 international admission
• Students 8 + 7 (2006 + 2007)
MBI programme organizers

Department of Computer Science, Department of Mathematics and Statistics, HY

Laboratory of Computer and Information Science, TKK

Faculty of Medicine, HY

Faculty of Biosciences, Faculty of Agriculture and Forestry, HY
Bioinformatics courses at the University of Helsinki

Department of Computer Science

- Practical course in biodatabases (II period): techniques for accessing and integrating data in biology databases.
- Biological sequence analysis (III period): basic probabilistic methods for modelling and analysis of biological sequences.
- Seminar: Regulatory networks (I & II periods)
- Seminar: Management of biological databases (III & IV periods)
Bioinformatics courses at the University of Helsinki

- Department of Mathematics and Statistics
  - Statistical methods in genetic epidemiology and gene mapping (I period)
  - Mathematical modelling (I & II periods)
  - Practical course on phylogenetic analysis (IV period): recommended to take also *Biological sequence analysis*
  - Adaptive dynamics (III & IV periods)
Bioinformatics courses at Helsinki University of Technology

Laboratory of Computer and Information Science

- Computational genomics (I & II periods): Algorithms and models for biological sequences and genomics
- Signal processing in neuroinformatics (I and II periods): overview of some of the main biomedical signal processing techniques
- High-throughput bioinformatics (III and IV periods): computational and statistical methods for analyzing modern high-throughput biological data
- Image analysis in neuroinformatics (III and IV periods): biomedical image processing techniques
Biology for methodological scientists (8 cr)

- Course organized by the Faculties of Bioscience and Medicine for the MBI programme
- Introduction to basic concepts of microarrays, medical genetics and developmental biology
- Book exam in I period (2 cr)
- Organized in three lectured modules, 2 cr each
- Each module has an individual registration so you can participate even if you missed the first module
- www.cs.helsinki.fi/mbi/courses/07-08/bfms/
Bioinformatics courses

- Visit the website of Master's Degree Programme in Bioinformatics for up-to-date course lists: http://www.cs.helsinki.fi/mbi
An introduction to bioinformatics
What is bioinformatics?

- Solving biological problems with computation?
- Collecting, storing and analysing biological data?
- Informatics - library science?
What is bioinformatics?

- Bioinformatics, n. The science of information and information flow in biological systems, esp. of the use of computational methods in genetics and genomics. (Oxford English Dictionary)

- "The mathematical, statistical and computing methods that aim to solve biological problems using DNA and amino acid sequences and related information."
  -- Fredj Tekaia

- "I do not think all biological computing is bioinformatics, e.g. mathematical modelling is not bioinformatics, even when connected with biology-related problems. In my opinion, bioinformatics has to do with management and the subsequent use of biological information, particular genetic information."
  -- Richard Durbin
What is not bioinformatics?

- Biologically-inspired computation, e.g., genetic algorithms and neural networks
- However, application of neural networks to solve some biological problem, could be called bioinformatics
- What about DNA computing?
Related concepts

- Computational biology
  - Application of computing to biology (broad definition)
  - Often used interchangeably with bioinformatics
- Biometry: the statistical analysis of biological data
- Biophysics: "an interdisciplinary field which applies techniques from the physical sciences to understanding biological structure and function" -- British Biophysical Society
- Mathematical biology “tackles biological problems, but the methods it uses to tackle them need not be numerical and need not be implemented in software or hardware.” -- Damian Counsell
Related concepts

- Systems biology
  - “biology of networks”
  - integrating different levels of information to understand how biological systems work
- Computational systems biology

Overview of metabolic pathways in KEGG database, www.genome.jp/kegg/
Why is bioinformatics important?

- New measurement techniques produce huge quantities of biological data
  - Advanced data analysis methods are needed to make sense of the data
  - Typical data sources produce noisy data with a lot of missing values
- Paradigm shift in biology to utilise bioinformatics in research
- To give you a glimpse of a typical situation in bioinformatics...
DNA microarray data

Outi Monni, Biochip Center, Biomedicum
Biological background

- Molecular Biology Primer: www.bioalgorithms.info