Algorithms for Computational Biology

Zsuzsanna Lipták

Masters in Molecular and Medical Biotechnology
a.a. 2015/16, fall term

Organisation

Organisation

- course times: Thu 10:30 - 12:30 (aula L), Fri 13:30 - 16:30 (aula H)
- language: English, but you can ask questions in English or in Italian
- webpage: coming soon!

  will include: current info, slides, materials etc.
- attendance: not obligatory but recommended
  (not all material will be available on webpage, plus:
  read what I wrote on the webpage of the course)

Organisation (cont.)

- email: `zsuzsanna.liptak@univr.it`
  Please put "Algorithms for Computational Biology" in the subject line
- office: CV 2, 1st floor, room 1.79
- student hours: Wed 10:30-12:30 **???**

Organisation (cont.)

- this course is 6 CFU
- exam:
  written and oral: admitted to oral only if you pass the written test
- for those who want this course to count for the Masters in Medical Bioinformatics (Dip. Inf.): it can replace 6 CFU of the "Fundamental Algorithms in Comp. Biology" **only if you do an additional assignment** on a topic from algorithms in computational biology (to be chosen together with lecturer)

Organisation (cont.)

- this course is 6 CFU
- exam:
  written and oral: admitted to oral only if you pass the written test
- for those who want this course to count for the Masters in Medical Bioinformatics (Dip. Inf.): it can replace 6 CFU of the "Fundamental Algorithms in Comp. Biology" **only if you do an additional assignment** on a topic from algorithms in computational biology (to be chosen together with lecturer)

Questions?

Goals of this course

1. to learn about some basic problems and algorithms behind common bioinformatics applications (sequence alignment, sequence similarity, phylogenetic reconstruction)
2. to get an idea of some basic computational issues involved (problem specification, efficiency, limitations)

Overview

- Part I: Sequence Analysis
  - Pairwise sequence alignment
  - **Detour:** Algorithm analysis
  - Multiple sequence alignment
  - String similarity and distance
  - Scoring matrices
  - Heuristic database search: FASTA, BLAST
- Part II: Phylogenetics
  - **Detour:** Trees and graphs
  - algorithms for distance-based data
  - character-based data, Perfect Phylogeny
  - Small Parsimony: Fitch's algorithm
  - Large Parsimony: heuristics

Books

- **Neil C. Jones and Pavel A. Pevzner**: An Introduction to Bioinformatics Algorithms (2004).—3 copies in library
- **David M. Mount**: Bioinformatics: Sequence and Genome Analysis (2004).—good book but very detailed!

Books

- **Neil C. Jones and Pavel A. Pevzner**: An Introduction to Bioinformatics Algorithms (2004).—3 copies in library
- **David M. Mount**: Bioinformatics: Sequence and Genome Analysis (2004).—good book but very detailed!
- **João Setubal, João Meidanis**: Introduction to Computational Molecular Biology (1997).—my favourite, 1 copy in library
- **Dan Gusfield**: Algorithms on Strings, Trees, and Sequences (1997).—the bible of string algorithms, a bit dated now
- **Joseph Felsenstein**: Inferring Phylogenies (2004).—important book on phylogenetics, very understandably written
- **Cormen, Leiserson, Rivest (& Stein)**: Introduction to Algorithms (different editions, 1990-onwards).—the bible of algorithms, a must have for anyone interested in algorithms (buy second hand, old editions are also fine)