Computational Methods for the Analysis of Biometric Data

CS 234

Welcome to CS 234

• Instructor: Stefano Lonardi
  – Office: WCH room 325
  – Phone: (951) 827-2203
  – Email: stelo@cs.ucr.edu

• Lectures: TR, 12:40-2:00pm INTS 2132

• Office hours: open door policy or by appointment

• http://www.cs.ucr.edu/~stelo/
  (click on “Teaching”, then CS 234 Winter16)
Course Format

• Three Homework
• One Project
• One Midterm
• One Presentation

The final grade is based on …

• Homework: 20%
• Midterm: 20% (first week March)
• Presentation (15 mins, in class, last two weeks) or Written Report (due finals’ week): 30%
• Project: 30% (due final week, ½ hour demo my office)
Course Overview

• Intro to Molecular Biology
• Intro to Probability and Statistics
  – Parameter estimation
  – Bernoulli and Markov models
  – Bayes Theorem
  – Entropy

Course Overview

• Pattern Modeling and Discovery (1D)
  – Deterministic patterns
  – Rigid patterns (Hamming distance)
  – Flexible patterns (Edit distance, Global alignment)
  – Position-specific matrix profiles (Relative entropy)
  – Hidden Markov Models (Viterbi, Baum-Welch)
  – Profile HMM
Course Overview

• Indexing and Searching Genomes (1D)
  – Short-read mapping
  – Hash Tables
  – Suffix trees
  – Suffix arrays
  – Burrows-Wheeler transform and FM-index
  – Software tools using the FM-index

Course Overview

• Biological networks (2D)
  – Co-expression networks
  – Gene regulatory networks
  – Protein-protein interaction networks (PPI)
  – Metabolic networks
  – Analysis of networks
  – Random models
What you will **not** learn here…

- Molecular biology (if you are really interested in Molecular biology, you should enroll/audit an undergraduate course in molecular biology, e.g. BIOL 107A/B)
- Statistics (same consideration … consider STAT160A-B, STAT161)
- How biological data is collected experimentally
- How to **use** existing software for molecular biology

If you want to know more about …

- Algorithms for multiple sequence alignment (local and global)
- Algorithms for fragment assembly
- Algorithms for restriction mapping and multiple digest mapping
- Algorithms for phylogenetic trees reconstruction
- …
… then you should also register “CS 238: Algorithms for Molecular Biology” by Prof. Jiang coming up next quarter

Some useful books
To learn about molecular biology

Lewin, Genes XI, Jones & Bartlett, 2012

Bioinformatics (sequences/probability)

Gusfield, Algorithms on Strings, Trees and Sequences, CUP, 1997
Durbin et al., Biological sequence analysis, CUP, 2003
Bioinformatics (references)

Krane, Raymer, Fundamental Concepts of Bioinformatics, Benjamin Cummings, 2003

Understanding Bioinformatics, Marketa Zvelebil, Jeremy O. Baum, Garland Science, 2007

Bioinformatics (references)

Jones, Pevzner, An Introduction to Bioinformatics Algorithms, MIT, 2004

Campbell, Heyer, Discovering Genomics, Proteomics and Bioinformatics, 2nd ed., Benjamin Cummings, 2006
Introduce yourself

• Name
• Department
• Graduate/undergraduate
• Years at UCR
• Research interests