CMSC423: Bioinformatic Algorithms, Databases and Tools

Fall 2008
Instructor: Mihai Pop
TuTh: 11-12:15, CSIC 1122
What is bioinformatics?

- Biology can be viewed as an information science (e.g. DNA is just a string of letters)
- Computers are essential in generating, managing, and analyzing biological data
- “Bioinformatics” or “Computational Biology” encompasses all applications of computers to the analysis of biological data
Why study bioinformatics?

• Exciting field! Help biologists figure out what life is all about.
• Work with people different from you – bio-geeks
• Many programmer/software engineer jobs in biotech industry currently filled by biologists – great need for people with CS backgrounds
Maryland's I-270 corridor, leapfrogging on federal agencies, is perhaps North America's fastest-growing source of industrial biotechnology (Science Magazine)

and many more...
Overview of course

• No knowledge of biology required
• Will cover areas of interest in CURRENT bioinformatics research
• Overall flow: data management (databases), data generation (sequencing), data analysis (extracting meaning)
• Examples based on real data (note: instructor spent 5 years in a biotech research institute)
Grading & workload

• Homework (10%)
  Goal: 4-5 assignments, each involving a couple hours of work at most
    – exercises from textbook
    – small programming assignments
    – “discovery” exercises (find something in public databases or using public software)

• Programming projects (15% + 15%)
  – Project 1 – assigned by instructor
  – Project 2 – chosen by student

• In-class midterms (25%) & final (35%)