BIOL/CS 383 - Bioinformatics

Spring, 2017
Chris Smith (crsmith.ant@gmail.com) and Charlie Peck (charliep@cs.earlham.edu)

Introduction

This course covers material from biology, computer science, mathematics, and from a magic space created by the synergy of the three. Bioinformatics is the application of mathematics and computer science to the field of biology. It offers much promise for the fields of disease origin, drug design, personalized medicine, public health, evolution, among many others. This course is a wide ranging introduction to the field, the tools, and the techniques used to work with large datasets, and will principally concentrate on the analysis of novel (i.e. never before seen) genomic and metagenomic data. The course will be centered around doing research and using tools, with much of the course time dedicated to active learning (by doing).

The prerequisites for Bioinformatics are BIOL 111, BIOL 112, or CS 128 (or consent of one of the instructors). If this doesn't describe you speak with Chris or Charlie RSN.

Materials

Practical Computing for Biologists (Haddock and Dunn 2010)
Journal articles, TED talks, Wikipedia pages, and other materials, as assigned.

Schedule

This is a 4 credit hour course which will meet for 1.5 hours 3 times each week. Class meetings will be a mix of assignment/project work, lecture, Q&A, dry and/or wet lab work. Class will meet on Tuesday at 8:30-9:50a, Wednesday at 2:30p-3:50p, and Thursday at 8:30-9:50a.

Mechanics

We'll use Moodle to organize descriptions of the readings, labs, projects, etc. and the assignments generally. Many of the underlying documents (e.g., lab and project descriptions) will be stored as Google Drive documents.

Tentative Organization and Goals (times are estimates)

Module 1: Introduction to DNA and code (3 weeks)
By the end of this module students will be able to:

1. do basic wet lab procedures and know the theory behind them (DNA extraction, gel electrophoresis, PCR)
2. use basic shell scripting
3. navigate DNA, RNA and amino acid sequence computationally

4. BLAST and custom genome searching

Module 2: Workflow, data, analysis (3 weeks)
By the end of this module students will be able to:

1. follow a standard operating procedure for a complex workflow
2. work on a remote computing cluster
3. describe what metagenomic data are and what they represent
4. use R to do community analyses

Module 3: Finding, downloading and analyzing large data sets (3 weeks)
By the end of this module students will be able to:

1. use an API to download data
2. integrate and build on analysis skills

Module 4: Dealing with a single genome, new tools and new analyses (3 weeks)
By the end of this module students will be able to:

1. work through layers of genome structure/organization
2. Map DNA/RNA reads to a genome (resequencing and RNAseq)
3. analyze data using a workflow (again)
4. Making sense of big datasets

Module 5: Using the resources and tools you've learned to make something (3 weeks)
By the end of this module students will be able to:

1. take a basic bioinformatics project from inception through to completion

Evaluation

1. Quizzes - 10% (These will cover all heard, read, and viewed material provided by us)
2. Homeworks and reports - 20%
3. Projects (module capstones) - 50%*
4. Attendance and participation - 20%

*Successful completion of each module capstone is a requirement for passing the course.

Academic Honesty

Often you will find it useful to discuss specific problems, techniques, etc. with fellow students. The sharing of ideas is a helpful and normal part of learning, and is encouraged. In particular one of the best ways to really learn something is to teach it to other people.

However, it's also possible to abuse those resources and turn-in work that isn't your own, particularly in computer science classes and when working in groups. See the Academic

Integrity Policy in the Curriculum Guide for the official Earlham College policy, http://www.earlham.edu/curriculumguide/academics/integrity.html

Disabilities

Please let one of us know as early in the semester as possible if there are any adaptations or accommodations you require, if there is any emergency medical information I should know about, or if you might need special arrangements in the case the building needs to be evacuated. The Earlham policy is:
Any student with a documented disability (e.g., physical, learning, psychiatric, vision, hearing, etc.) who needs to arrange reasonable accommodations must contact Academic Support Services and the instructor at the beginning of each semester. Accommodation arrangements must be made during the first-two weeks of the semester.

It is important to follow this procedure.

Standard Disclaimers and Advice

Showing-up every time is 50% of success, bringing “game” is another 25%. Perhaps natural skill, talent, and serendipity account for the remainder.