Computer Science & Engineering Syllabi   College of Engineering & Computer Science

Fall 2005

CS/BIO 471/671: Algorithms for Bioinformatics

Michael L. Raymer
Wright State University - Main Campus, michael.raymer@wright.edu

CS/BIO 471/671: Algorithms for Bioinformatics
Fall, 2005

Course Description

Theory-oriented approach to the application of contemporary algorithms to bioinformatics. Graph theory, complexity theory, dynamic programming and optimization techniques are introduced in the context of application toward solving specific computational problems in molecular genetics. 4 credit hours.

Meeting Time and Place

2:45 – 4:00 pm, Monday and Wednesday
155 Russ Center

Text Books and Other Materials

- Required text:

    - J. Setubal and J. Meidanis, **Introduction to Computational Molecular Biology**.

- Reference texts:

    - D. E. Krane and M. L. Raymer **Fundamental Concepts in Bioinformatics**.
    - Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, Clifford Stein, **Introduction to Algorithms**, 2nd edition.

Instructors and Office Hours

Dr. Michael L. Raymer
338 Russ Engineering Center
775-5110

michael.raymer@wright.edu
http://www.wright.edu/~michael.raymer
Tue/Thurs; 3:00 – 5:00 pm

Dr. Dan Krane
126 BH
775-2257

dan.krane@wright.edu
Mon/Wed; 12:30 – 2:30 pm

Course Web Page

The course web page will be the primary method for distributing course announcements, supplemental materials, class notes, etc. Please check the page often. The URL is:
    http://birg.cs.wright.edu/cs471

Grading

Course grades will be computed based on the following elements:

10%	Homework
30%	Midterm Exam I
30%	Midterm Exam II
30%	Term Project

Final grades will be based on the standard university-wide score divisions (i.e. 90%, 80%, 70%, etc.). However, the instructors *may* curve the final grades if such action is merited by the final score distribution for the course.

Course Content & *Tentative* Lecture Schedule

Week 1: Basic concepts of molecular biology – *Chapter 1 of Krane & Raymer text, and Chapter 1 of Setubal & Meidanis*
- The mechanisms of molecular genetics, transcription, translation, and protein synthesis
- Junk DNA, reading frames and chromosomes, repeated DNA
- Is the genome like a computer program?
- How the genome is studied

Week 2: Data structures and complexity – *Chapters 1 & 2 of Cormen, Leiserson, Rivest*
- The RAM model of computing
- Time and space complexity
- Design and analysis of simple sorting algorithms
- Asypmtotic notation
- Counting methods and combinatorics
- Polynomial and non-polynomial growth, introduction to NP-completeness

Week 3: Repeated DNA sequences, mutation, and evolution
- Types and sources of repeated DNA
- Sources of error: Failures in DNA replication and repair
- Experimental design and group organization for term project
- Tools for working with sequences

Week 4: Sequence comparison and searching – *Chapter 3 of Setubal & Meidanis*
- Global, local, and semi-global comparison
- General and affine gap penalty functions
- Comparing multiple sequences
- BLAST and FASTA
- Sequence searching statistics and significance

Week 5: Introduction to Graph Theory and Optimization Algorithms – *Section 2.2 of Setubal & Meidanis*
- Introduction to graph terminology and techniques
- Greedy strategies, exhaustive strategies, and dynamic programming
- 0-1 knapsack problem

Week 6: Fragment assembly of DNA – *Chapter 4 of Setubal & Meidanis*
- The biological process of DNA sequencing
- The need for fragment assembly
- Models for fragment assembly: Shortest common superstring, reconstruction, and multicontig.
- Biological complications: base call errors, unknown orientation, repeated regions, etc.

Week 7: Physical mapping of DNA – *Chapter 5 of Setubal & Meidanis*
- Restriction site mapping, hybridization mapping
- Mathematical models for mapping strategies

Week 8: Phylogenetic trees
- Trees and distances
- Distance-based approaches
- Character-based approaches
- Applications: PAM and BLOSSUM

Weeks 9 & 10: Tools and Techniques for Bioinformatics – *Chapter 8 of Setubal & Meidanis*
- The protein folding problem
- Branch and bound techniques for protein threading
- Genetic Algorithms
- Classifiers: Nearest neighbor and Bayesian methods
- Neural networks

Week 10: Last Class
- Class presentations of term project results

Policies & Notes

Homework Assignments are due by the end of the lecture period on the assigned due date. Late homeworks cannot be accepted.

Collaboration: Discussion of course contents with other students is an important part of the academic process and is encouraged. However, it is expected that course programming assignments, homeworks, and other course assignments will be completed *on an individual basis* unless the assignment states otherwise.

Students may discuss general concepts with one another, but may not, under any circumstances, work together on the actual implementation of any course assignment. **If you work with other student on "general concepts" be certain to** *acknowledge* **the collaboration and its extent in the assignment. Unacknowledged collaboration will be considered dishonest.** "Code sharing" (including code from previous quarters) is strictly disallowed. "Copying" or significant collaboration on any graded assignments will be considered a violation of the university guidelines for academic honesty.

If the same work is turned in by two or more students, all parties involved will be held equally accountable for violation of academic integrity. *You are responsible for ensuring that other students do not have access to your work*: do not give another student access to your account, do not leave printouts in the recycling bin, pick up your printouts promptly, do not leave your workstation unattended, etc. If you suspect that your work has been compromised notify an instructor immediately.

Note: Failure to attend the first day of class, during which time these academic honesty policies will be explained in detail, *does not excuse you from following these policies*. If you have any questions about collaboration or any other issues related to academic integrity, please see one of the course instructors immediately for clarification.