SPECIAL TOPICS:
Computational Evolutionary Biology

Class Meeting

Lectures:
Mondays and Wednesdays 0905-1015 Dirac Science Library Room 156
Lab-session:
Mondays 1100-0100 Dirac Science Library Room 156

Instructors

Fredrik Ronquist	Peter Beerli
150-S DSL	150-T DSL
ronquist@csit.fsu.edu	beerli@csit.fsu.edu
645-1325	645 1324

Office Hours

Monday 0100-0300 or by appointment.

Objectives

This course will introduce students to methods used in phylogenetics and population genetics and writing computer programs using such methods. Primary objectives of the course are:

1. to expose students to a large set of modern methods used in the field of theoretical evolutionary biology

2. introduce students to the programming aspects of the field

3. enable students to develop software that is a base for further exploration or incorporation of new methods.

Content

Advanced computational methods are becoming increasingly important in biology. A wide range of applications — including, for instance, identifying pathogens, tracing viral transmission pathways, and reconstructing the geographic expansion of humans out of Africa — rely on evolutionary inference. This course will cover the methods currently used for evolutionary inference, the stochastic models and inference principles they are based on, and how they are implemented in practice. In a separate lab session, the students will get hands-on experience in developing computational software implementing these methods. We expect that the students leave the course with the necessary skills to develop their own ideas and methods.

Textbook

Felsenstein, Joseph. 2004. Inferring Phylogenies. Sinauer, Sunderland MA (we did not preorder the books through the bookstore, we recommend ordering it through amazon.com, much faster than ordering directly from Sinauer)

The textbook includes material for almost all of the major topics that we expect to cover. There is a small amount of material that will be presented in class that is not covered by the textbook. This textbook is a great starting point for further exploration of the subject and we certainly will not have time to explore all chapters.

Grading

- Grades will be based on students' execution of the 10 programming assignments, each of which involves understanding the algorithms, code design, and program documentation (Points for working code and beauty of code (OOD principles, clarity in comments) [5 points each]

- Each student will do a project on their own during the last 4 weeks of the semester and also give a short presentation of their work in the last regular lab-meeting. [20+10 points]

- We will have a theory test on October 19th (midterm) and a final test in finals week. [10 points, each]

A student who accumulates 90% or more of the possible points will receive a grade of "A", a student who accumulates between 80% and 89% of the possible points will receive a grade of "B", a student who accumulates between 70% and 79% of the possible points will receive a grade of "C", a student who accumulates between 60% and 69% of the possible points will receive a grade of "D", and a student who accumulates less than 60% of the possible points will receive a grade of "F".

Attendance

Attendance in class is strongly recommended. While there is no penalty for failure to attend, the instructor may not be able to review with you any material that you miss by nonattendance unless you have a valid reason for absence (illness, death of a family member, professional conflict, subpoena to court, varsity athletics, religious holiday) or make a prior arrangement.

Missed/Late Assignments

Deadlines for assignments will be announced in class; late assignments will be accepted for full grade only in cases of illness or death in the family. 50% of the points are deducted for late assignments.

Florida State Honor Code

You are expected to know and adhere to the FSU Honor Code.

ADA Policy

If you need alternative accommodations for any reason, please notify the instructor.

Lectures

1. Introduction. Trees and tree representation (Aug. 29; Fred)

2. Parsimony. Counting evolutionary change (Aug 31; Fred)

    - Fitch parsimony
    - Wagner parsimony
    - Other parsimony variants
    - Sankoff cost matrices

3. Searching for the best tree(s) (Sept. 7; Peter)

    - How many trees are there
    - Explicit enumeration
    - Branch and bound
    - Heuristic methods

4. Stochastic models of evolution I (Sept. 12; Peter)

    - Nucleotide models: JC, F81, HKY, GTR
    - Rate variation across sites: Gamma

5. Stochastic models of evolution II (Sept. 14; Fred)

    - Protein models
    - Codon models
    - Doublet models

6. Stochastic models of evolution III (Sept. 19; Peter)

    - Restriction sites
    - Microsats
    - AFLP

7. Stochastic models of evolution IV (Sept. 21; Fred)

    - Quantitative characters
    - Morphology

8. Hidden Markov models (Sept. 26; Fred)

    - General introduction to
    - Rate variation across sites: the autocorrelated gamma
    - Mixture models

9. Maximum likelihood inference I (Sept. 28; Fred)

- General principles
- Coin-tossing
- Statistical consistency and efficiency

10. Maximum likelihood inference II (Oct. 3; Peter)

- Conditional likelihoods
- Calculating likelihoods of evolutionary models

11. Bayesian inference (Oct. 5; Fred)

- General principles
- Coin tossing
- Statistical consistency and efficiency
- Priors

12. Markov chain Monte Carlo I (Oct. 10; Peter)

- Gibbs and Metropolis samplers
- Calculating Hastings ratios

13. Markov chain Monte Carlo II (Oct. 12; Fred)

- Convergence and mixing
- Metropolis-coupling

14. Review (Oct. 17; Peter and Fred)

15. Mid-term exam (Oct. 19; students)

16. Trees and tree models I (Oct. 24; Fred)

- Molecular clocks
- Relaxed clock models

17. Trees and tree models II (Oct. 26; Fred)

- Birth-death process
- Covarion models

18. Population models (Oct. 31; Peter)

19. The coalescent I (Nov. 2; Peter)

- The simple n-coalescent
- Relation to other population genetic models

20. The coalescent II (Nov. 7; Peter)

- The structured coalescent
- Coalescent and selection and other complications

21. Implementing the coalescent (Nov. 9; Peter)

- ML implementations
- Bayesian implementations

22. Phylogeography (Nov. 14; Peter)

- Haplotype networks
- Nested clade analysis

23. Historical biogeography (Nov. 16; Fred)

- General overview of problems
- Parsimony reconstruction of vicariance scenarios
- Parsimony study of dispersal events
- Statistical approaches

24. Gene trees and species trees; hosts and parasites (Nov. 21; Fred)

- Overview of the field
- Parsimony methods
- Statistical methods

25. Model selection and model averaging I (Nov. 23; Johan)

- Hierarchical and nonhierarchical models
- Hierarchical likelihood ratio test
- Akaike information criterion
- Bayes information criterion

26. Model selection and model averaging II (Nov. 28; Johan)

- The Bayesian approach: Bayes factors
- Model averaging

27. Bootstrapping and jackknifing (Nov. 30; Peter)

- Nonparametric and parametric approaches to confidence
- Phylogenetic implementations
- Bootstrap corrections
- Bootstraps and posterior probabilities

28. Statistical multiple sequence alignment (Dec. 5; Peter)

- Heuristic dynamic programming approaches
- Bayesian MCMC sampling of multiple alignments

29. Review (Dec. 7; Peter and Fred)

30. Final exam (Dec. 12-16; students)

Lab-sessions

1. Introduction to Java and NetBeans. Parsing and printing trees (one week plus Labor Day week) (Aug. 29; Fred)

2. Calculating the parsimony score of a tree (one week) (Sept. 12; Fred)

3. Searching for the best tree (Sept. 19; Peter)

    - Exhaustive search
    - Branch and bound

4. Searching for the best tree (heuristic) (Sept. 26 Fred)

    - Stepwise
    - Branch rearrangement

5. Simulate data on a tree (one week) (Oct. 3; Peter)

6. Calculate likelihood of a tree (one week) (Oct. 10; Fred)

7. Optimization of likelihood (one week) (Oct. 17; Peter)

8. MCMC sampling of phylogenetic models (one week) (Oct. 24; Fred)

9. Simulate a coalescent tree (one week) (Oct. 31; Peter)

10. Model selection using Bayes factors (one week) (Nov. 7; Fred)

11. Individual project (four weeks) (Nov. 14; Peter and Fred)

12. Oral presentation of individual project (Dec. 5; students)