

GENERAL COLLOQUIUM
SEMINAR

Enabling Highly Accurate Large-Scale Phylogenetic Estimation

Shel Swenson
University of Southern California

Abstract: Evolutionary histories of sets of molecular sequences are a fundamental tool in many biological and biomedical questions of societal importance, including biodiversity conservation, drug development, and even forensic investigations. The best methods for estimating these evolutionary histories, or phylogenetic trees, are based on NP-hard optimization problems, and thus phylogenetic analyses of large-scale datasets is extremely computationally intensive. The continually diminishing costs and increasing throughput of DNA sequencing technologies will lead to an ever greater demand for methods capable of producing accurate phylogenetic trees on complex, large-scale molecular datasets. In this talk, I will describe algorithms my collaborators and I have developed to address this demand.

I will present SuperFine and ASTRAL, two divide-and-conquer approaches with desirable theoretical properties and excellent empirical performance. Both methods are supertree approaches in that they divide a larger taxon set into subsets, estimate trees on those subsets, and apply a supertree method which assembles a tree on the entire set of taxa from the smaller "source" trees. SuperFine is designed to handle datasets with source tree conflict only due to estimation error, while ASTRAL is designed to handle source tree conflict due to estimation error and incomplete lineage sorting which can cause gene trees to differ from the underlying species tree.

I will present supertree methods in a mathematical context, focusing on some theoretical properties of MRP (Matrix Representation with Parsimony), the most popular supertree method, and SuperFine which outperforms MRP. I will also describe a desirable statistical property of ASTRAL and this method's potential to enable highly accurate genome-scale phylogenetic analysis.

Tuesday, March 25, 2014, 4:00 pm
Mathematics and Science Center: W303

MATHEMATICS AND COMPUTER SCIENCE
EMORY UNIVERSITY