
Discrete and algorithmic mathematics in biology and epidemiology - Part II

(Org: Pengyu Liu (Simon Fraser University))

JULIA PALACIOS, Stanford University

Distance-based summaries and modeling of evolutionary trees

Ranked tree shapes are used to model evolutionary relationships for evolutionary biology and phylodynamics. Bayesian methods explore the posterior distribution of trees, however assessing uncertainty and summarizing distributions remains challenging. Similarly, in many instances, one seeks to summarize samples of trees obtained from different samples and environments, and wishes to assess stability and generalizability of these summaries. Here, we exploit recently proposed distance of unlabeled evolutionary trees and provide a combinatorial optimization algorithm for estimating Fréchet means and variances. We show the applicability of our summary statistics for studying popular tree distributions and for studying the evolution of viruses.

CHRISTOPH WEITKAMP, Universität Göttingen

GROMOV-WASSERSTEIN BASED PHYLOGENETIC TREE SHAPE COMPARISON

The shapes of phylogenetic trees carry important information about short and long-term evolutionary processes. In this talk, we propose to model the shapes considered as dissimilarity measure spaces and to compare them based on the Gromov-Wasserstein distance. We quantify their similarity on the basis of a dissimilarity-preserving soft assignment of the elements of the respective spaces. To reduce the computational complexity, we introduce several computationally efficient surrogates and illustrate in various examples that these represent powerful tools for the comparison of phylogenetic tree shapes. We compare our methods to the Colijn-Plazotta metric and illustrate the fundamental differences between these approaches.

JIANRONG YANG, Sun Yat-sen University

Developmental cell lineage trees, and the quantitative comparisons between them

The developmental process of multicellular organisms can be summarized as the cell lineage tree (CLT). Technological breakthroughs have facilitated determination of more CLTs, but complete comprehension of the data remains difficult without a quantitative framework for CLT comparison. I will briefly introduce some CLT data, and present "DELTA", a novel algorithm that quantitatively compares CLTs. Application of DELTA to the worm CLTs allowed inference of transcriptomic resemblance, identifying cell fate transformations, predicting functional relationships between mutants, and finding evolutionary correspondence between cell types of different species. Quantitative comparison between CLTs can likely help answer many questions surrounding the developmental process.

LOUXIN ZHANG, National University of Singapore

The Bourque Distances for Mutation Trees of Cancers

Mutation trees are rooted trees in which nodes of arbitrary degree are labeled with a mutation set. These trees are used in computational oncology to represent the mutational history of tumours. Classical tree metrics are of limited use for the comparison of mutation trees. One reason is that mutation trees inferred with different methods or for different patients often contain different sets of mutation labels. We generalize the Robinson-Foulds distance into a set of distance metrics called Bourque distances for comparing mutation trees.