

Formal power series and combinatorial methods in phylogenetics

Mike Steel (University of Canterbury)

Phylogenetics is the reconstruction and analysis of evolutionary trees in systematic biology and other areas of classification (e.g. historical linguistics, epidemiology). The mathematics that underlies this field is based on combinatorics and discrete random processes. In this talk, I will highlight both established and recent phylogenetic applications of familiar combinatorial techniques that have proved useful for deriving new results on phylogenetic trees. In particular, I will describe how:

- exponential generating functions for trees and forests can be used (together with Menger's theorem and multivariate Lagrange inversion) to enumerate phylogenies under a 'minimal evolution' score;
- every binary phylogenetic tree can be realized as a unique 'perfect phylogeny' with just four functions ('characters') from the leaf set of the tree into an infinite discrete state space;
- the probabilistic method provides an $O(n^\alpha)$ (for any $0 < \alpha < 1$) analog of the last result when the state space is finite and the tree has n leaves.
- extended Pólya urn models are relevant to speciation-extinction models that 'evolve' phylogenetic trees.