
On trees, tanglegrams, and tangled chains

Sara Billey^{*1}, Matjaz Konvalinka^{*2}, and Frederick Matsen Iv^{*3}

¹Department of Mathematics, University of Washington, USA – United States

²Department of Mathematics, University of Ljubljana Institute for Mathematics, Physics and Mechanics – Ljubljana, Slovenia

³Computational Biology Program, Fred Hutchinson Cancer Research Center – Seattle, WA 98109, United States

Abstract

Tanglegrams are a class of graphs arising in computer science and in biological research on cospeciation and coevolution. They are formed by identifying the leaves of two rooted binary trees. The embedding of the trees in the plane is irrelevant for this application. We give an explicit formula to count the number of distinct binary rooted tanglegrams with n matched leaves, along with a simple asymptotic formula and an algorithm for choosing a tanglegram uniformly at random. The enumeration formula is then extended to count the number of tangled chains of binary trees of any length. This work gives a new formula for the number of binary trees with n leaves. Several open problems and conjectures are included along with pointers to several followup articles that have already appeared.

*Speaker