Patching Up X-Trees

Sebastian Böcker[1], Andreas W.M. Dress[1], and Mike A. Steel[2]

[1] GK Strukturbildungsprozesse, FSP Mathematisierung, Universität Bielefeld, PF 100 131, 33501 Bielefeld, Germany {boecker,dress}@mathematik.uni-bielefeld.de

[2] Biomathematics Research Centre, University of Canterbury, Private Bag 4800, Christchurch, New Zealand m.steel@math.canterbury.ac.nz

Abstract. Trees are widely used to represent evolutionary, historical, or hierarchical relationships in various fields of classification. In biology, such trees ("phylogenies") typically represent the evolutionary history of a collection of extant species, or the line of descent of some gene. In most applications, the objects of interest are found at the tips (leaves) of the tree, while all interior vertices correspond to a branching (or speciation) event.

In our contribution, we will present a surprisingly explicit solution to a particular problem in tree reconstruction dealing with the formation of "supertrees" from a family of subtrees.

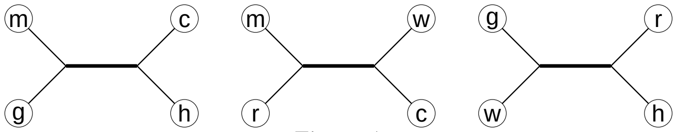

Figure 1

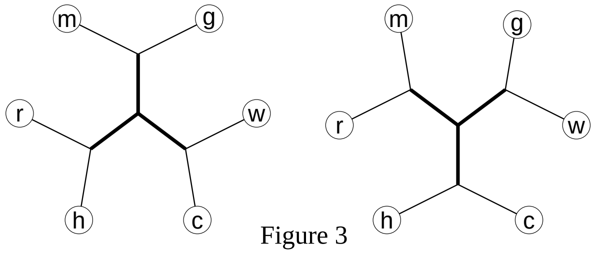

Figure 3

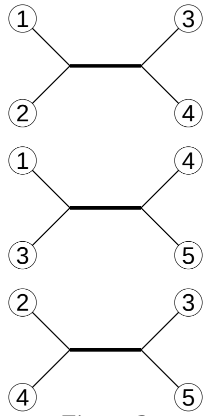

Figure 2

To explain this problem, consider (cf. Fig. 1) the three (unrooted) "quartet trees" mg|ch, mr|wc, and gw|rh, each relating just four out of the altogether six species man (m), gibbon (g), cow (c), whale (w), horse (h), and rhino (r); or (cf. Fig. 2) the three quartet trees 12|34, 13|45, and 24|35. In each case, one desires to determine a (hopefully unique) supertree "displaying" all of these three quartet trees simultaneously. Remarkably, while there is not even one supertree for the three quartet trees from Fig. 2, there are exactly two possible supertrees for the first three quartet trees (cf. Fig. 3) — a fact that, by the way, implies a simple, though fundamental restriction regarding the existence of "unbiased" consensus and supertree methods (cf. [4]).

So, the question arises to find methods to automatically generate the set of all supertrees that display a given family of partial trees. In the poster, we will

1. report a result (cf. [2]) that, for any collection X of n distinct species, completely characterizes those families of *exactly* $n-3$ quartet trees (with leaf labels drawn from X) that allow for exactly one supertree;
2. demonstrate, by means of examples, some remarkable and decisively non-trivial consequences of our result;
3. detail algorithmic applications; and
4. point out how our method can be used to explore the (possibly empty) set of all supertrees that display a given family of quartet trees.

References

[1] Sebastian Böcker and Andreas W.M. Dress, *Patchworks*, to appear in Adv. in Math., 1999.

[2] Sebastian Böcker, Andreas W.M. Dress, and Mike Steel, *Patching up X-trees*, Ann. Comb. **3** (1999), 1–12.

[3] Péter L. Erdős, Michael A. Steel, László A. Székely, and Tandy J. Warnow, *A few logs suffice to build (almost) all trees. I*, Random Structures Algorithms **14** (1999), no. 2, 153–184.

[4] Mike Steel, Andreas W.M. Dress, and Sebastian Böcker, *On the non-existence of of consensus methods that are equivariant and Pareto on subtrees*, submitted.