

# MATHEMATICAL PHYLOGENETICS AND THE SPACE OF TREES

{ MERCEDES COLEMAN, CODY FITZGERALD, AMBER HOLMES, EMILY SMITH }  
GRADUATE ASSISTANT: COLBY LONG AND MENTOR: SETH SULLIVANT.



## INTRODUCTION

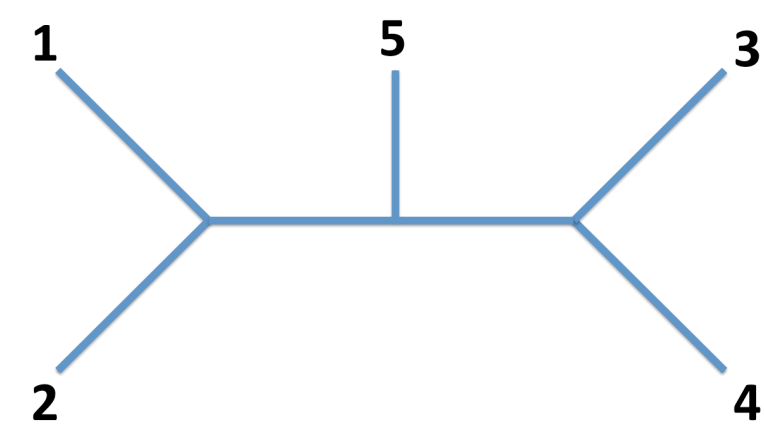
Phylogeny is the branch of biology that focuses on illustrating and studying the relationships between a set of species. Mathematical phylogenetics studies the structure and properties of these relationships which we represent as trees. Using ideas from graph theory, we discuss the space of all possible trees and how to traverse a path between two trees in this space. We can also compute the average of a set of trees in a variety of metrics; these mean trees are interesting because variations in data from the same species or even different software packages using the same data often build different trees to represent one set of relationships. It is then useful to examine multiple averages to determine the most representative tree for a particular set of relationships. Our project focuses on implementing algorithms for these different mean trees in MATLAB and developing an algorithm for the  $(1,\infty)$ -mean tree.

## SPLITS EQUIVALENCE THEOREM

- An **X-split** is a partition of  $X$  into two non-empty sets. We denote the X-split whose blocks are  $A$  and  $B$  by  $A|B$ .
- A pair of X-splits  $A_1 | B_1$  and  $A_2 | B_2$  are **compatible** if at least one of the sets  $A_1 \cap A_2$ ,  $A_1 \cap B_2$ ,  $B_1 \cap A_2$ , and  $B_1 \cap B_2$  is the empty set.

**Theorem 1 Splits-Equivalence Theorem.** Let  $\sum$  be a collection of X-splits. Then, there is a tree,  $T$ , such that  $\sum = \sum(T)$  if and only if the splits in  $\sum$  are pairwise compatible. Moreover, if such a tree exists, then, up to isomorphism,  $T$  is unique.

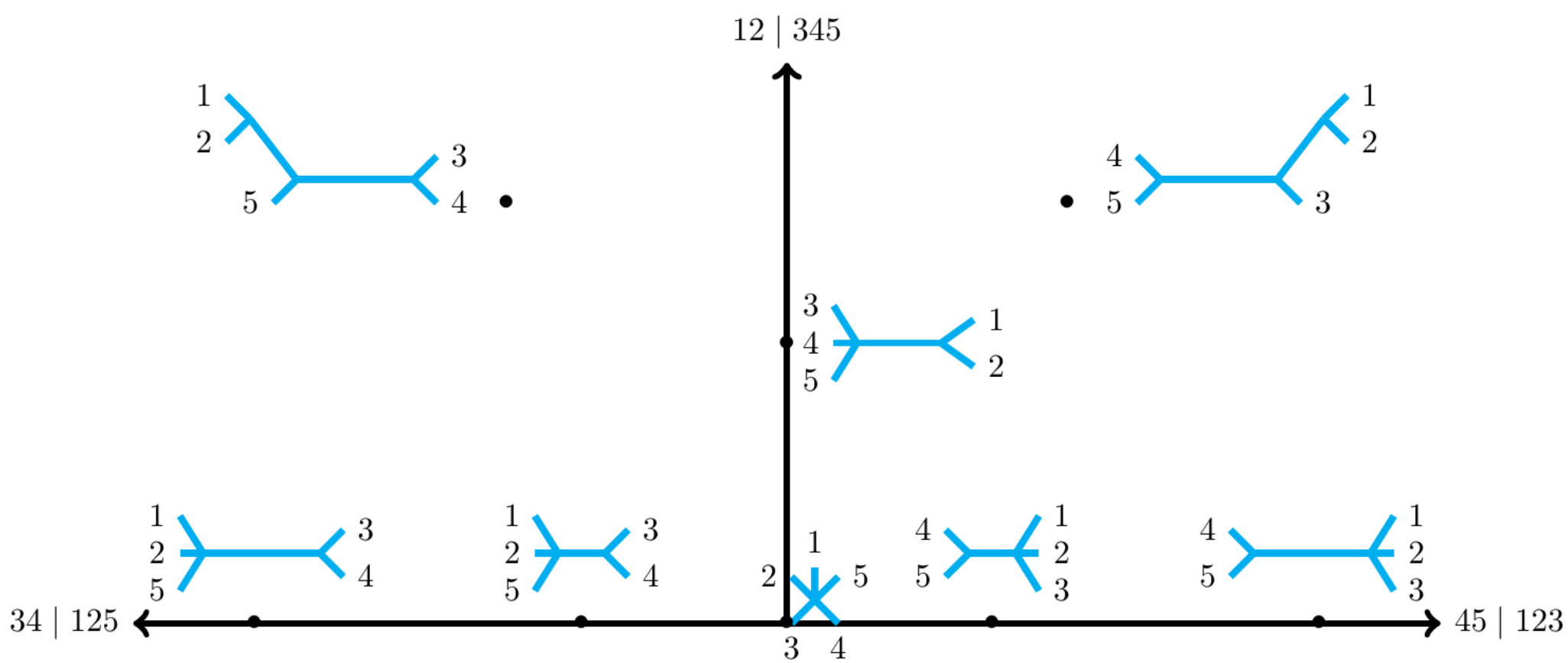
1		2345
2		1345
3		1245
4		1235
5		1234
12		345
34		125



## TREE SPACE

Tree space for n-leaf trees:

- Tree space** is a space constructed such that every point represents a possible tree.
- Non-trivial splits determine rays to form orthants.
- Orthants are "glued" together along common boundary rays.
- Distance away from the origin represents edge length.



## REFERENCES

1. Semple and Steel, *Phylogenetics*.
2. Billera, Louis J., Holmes, Susan P., Vogtmann, Karen. Geometry of the space of phylogenetic trees.
3. Megan Owen, Scott Provan. A Fast Algorithm for Computing Geodesic Distances in Tree Space.
4. Ezra Miller, Megan Owen, J. Scott Provan. Polyhedral computational geometry for averaging metric phylogenetic trees.

## GEODESIC IN TREE SPACE

Geodesic:

- The unique shortest path from one point to another.
- Exists for any two points in tree space because the space is CAT(0).

