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 ways; 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 set of relationships. Our project focuses on implementing algorithms for these different mean trees in MATLAB and developing an algorithm for the (1,∞)-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 A1:B1 and A2:B2 are compatible if at least one of the sets A1 ∩ B2, A1 ∩ B1, A2 ∩ B2, and B1 ∩ B2 is the empty set.

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

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).

For any two trees, one tree (T1) will have a set of edges A and the other tree (T2) will have a set of edges B. These sets can be partitioned into disjoint sets of splits denoted A1, A2, ..., Aτ, from T1 and B1, B2, ..., Bτ, from T2, such that the sequence of orthants traversed from T1 to T2:

A = (c1, c2, c3) = A1 − (c2, c3)
B = (c4, c5, c6) = A2 − (c3, c6)
A3 = (c4, c5)
B3 = (c6)

(2,2)-mean trees
(2,2)-mean trees are found using Sturm’s algorithm, an iterative algorithm that takes the form

M1 = T1
M2 = 1/2 M1 + 1/2 T2

... for all n ≥ 2

Mn = k−1/2 Mn−1 + 1/2 Tn

Let M(∞) = limn→∞ Mn denote the point in L∞ that is λn dL∞ perpendicularly above the line segment between x and y.

MATLAB Implementation

(2,2) mean:
- A set of trees and the desired number of iterations is input. Approximations are computed as described by Sturm’s algorithm (Ref. 4) with help of geodesic and intermediate tree calculation described in Ref. 3. The final approximation is output as the mean tree.

(1,∞) mean:
- A set of trees is input. The majority consensus tree (Ref. 1) is determined because these splits are the only ones with nonzero edge length in the mean tree. The median edge length of each split is computed, and the majority consensus tree along with these edge lengths is output as the mean tree of the L∞ geodesic path.

Note: Trees are generally stored in Newick format, in which a semicolon concludes a tree. The following string represents one tree:

```
(1,2,3,4,5);6;7,8;9
```

References
1. Semple and Steel, Phylogenetics.

Acknowledgements
- Funded by: NASA and NSF grants.
- Supported by: North Carolina State University.

Contact Information
- M. Coleman mcoleman@lamar.edu.
- C. FitzGerald ccf29@wlu.edu.
- A. Holmes abhyolmes@student.lagrange.edu.
- E. Smith smithel3@kennesyon.edu.