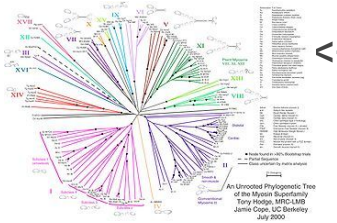
# Navigating Phylogenetic Trees using Graphing Algorithms

Giorlando Ramirez

# So...What's the problem?

-Phylogenetic trees can be huge, and calculating the distance in between species that are far away from each other in the tree can be difficult.




<- Big tree

-This makes it difficult to find the shortest evolutionary distance(branch lengths) in between two species.

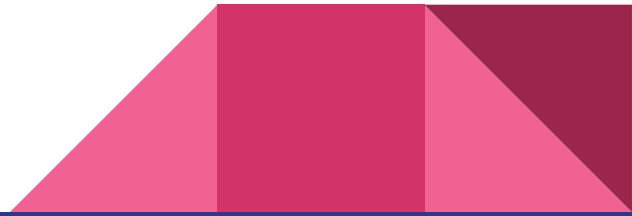
# The solution to our problems?

**Breadth-first Search***

**Input:** A Graph, a starting node in the Graph and goal node in the Graph.

**Output:** Goal node. The shortest path back to starting node.

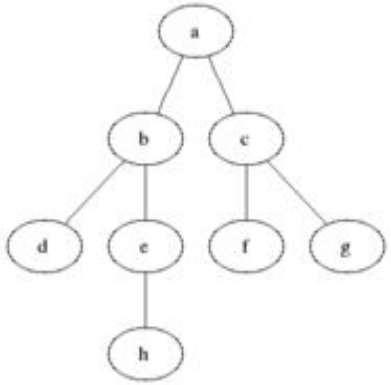
*Making this function was hard.



# So how does it work?

-Visits neighbor nodes first before moving on to next level neighbors. As opposed to simply going the full depth(Depth first search).

Here’s a nice gif to demonstrate exactly how it works:




Gif source: https://en.wikipedia.org/wiki/Breadth-first_search

# Data and Formatting

-Works on data sets presented in the Newick format:

Example:
(Bovine:0.69395,(Gibbon:0.36079,(Orangutan:0.33636,(Gorilla:0.17147,(Chimp:0.19268, Human:0.11927):0.08386):0.06124):0.15057):0.54939,Mouse:1.21460):0.10

-Format on Python:

Dataset1 = [['Chimp','A'],['Human','A']]

Dataset1Dictionary = {('Chimp','A'):0.19268,('Human','A'):0.11927}

# Results

-The program can:

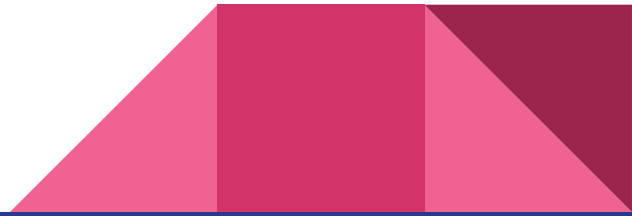
-Trace the path taken between species.

-Calculate number of edges.

-Find the distance between species.

-Good news: Program works on practice Data Set. And creates nice graph in GraphSpace.

-Bad news: Wasn’t able to find a Newick Tree format database to download loads of data directly from. So Data-wise the program is very limited. And I had to build all of the test graphs and corresponding dictionaries.



# Now, the program.

I will now present my extremely underwhelming somewhat interactive program.