CPR/SE 491 Weekly Report   MAY 15-10   Week 5

Advisors: Ruchi Chaudhary
Client: Gordon Burleigh
Members (roles):   Ben   -- Team leader
Cole   -- Team Key Concept Holder
Ian   -- Webmaster
Paul   -- communication leader

Project Title: Gene Tree Improvement Tool / TreeFix

Weekly Summary

This has been our most productive week yet. With Ruchi, we discussed how the RF cost algorithm works. Here, we learned that the easiest way to implement it was via an algorithm called LCA, which calculates relevant information about the non-root internal nodes of a tree in constant time. We then performed a breadth-first analysis of the code as a group. Through this research, we were able to identify specific functions and classes that will have to be evaluated and modified in order to accept the new cost method to be introduced. Our greatest achievement this week in terms of work eliminated, was identifying that the code would not need to be deeply refactored in order to support non-binary trees. A library file treelib.py was inspected, and certain patterns about the data structures contained within implied that it offered the support for non-binary trees we needed.

 Paul - analyzed the "Search_landscape" function, which deals with choosing a correct minimum cost tree given a specified distance method
 Ian - analyzed a library file "phylo.py." Specifically, the function **recon_root**, and the functions in that same library which recon_root depends on. recon_root symbolizes a very important crux of our project, which is the successful implementation of a cost reconciliation model. It was uncovered that this function recon_root is specific to the dup-loss method of cost reconciliation, and therefore incompatible with MulRF.
Ben - analyzed the main method in treefix.py. Went through and catalogued all external calls, and which packages contain those calls
Cole - analyzed the parse tree function to see if the current code can allow non-binary tree inputs, and catalogued what sections need to be refactored.

Meeting notes:

10/3   Group Meeting with Advisor
Duration: *40 min*   **Members Present:** *All*

Purpose and Goals:
Go over the two papers Ruchi provided us on dupLoss and RF mapping
Learn the algorithms behind LCA mapping, as well as how RF distance is calculated
Achievements:
Received clarification about RF cost calculation algorithm.
Learned methodology of LCA Mapping, which is used in calculating RF, and dupLoss algorithms

10/3   Group Meeting
Duration: *4.5 hr*   **Members Present:** *All*

Purpose and Goals:
Finished the first project plan draft and weekly report.
Examined code for TreeFix.

Achievements:
Identified some areas that will have to be modified in order to implement the RF cost method.
Achieved a greater understanding of the inner workings of TreeFix(not 100% but ~80%)
Individual achievements are in the weekly summary

Pending issues
1. Review TreeFix code to determine what code can be reused.
2. Develop a format for a general cost reconciliation algorithm

Plans for next week
Paul - Look at existing RF algorithm code to determine if we can reuse any of it for Treefix. Determine which parts of the code need to be changed to be able to accommodate RF algorithm.
Ben - understand the returns from the LCA algorithms and hope they can be used in the RF cost design
Cole - read through the dupLoss algorithm in phylo.lib to see what calls are made to it, and where we need to implement our RF algorithm.
Ian - Research into the phylo.py recon_root() method and if it is only useful towards the dup-loss cost method

Individual Contributions(this week)
Ben Streit(6.5)
Ian Ray (8hr)
Paul Leichty (6.5hr)
Cole Poffenberger (6 hr)

Total contributions for the project
Ben Streit(11.75)
Ian Ray(13)
Paul Leichty (10.5)
Cole Poffenberger(10.75)