

A PROPOSAL TO DO RESEARCH ON VISUALIZATION OF A PHYLOGENETIC TREE

By Kathleen Timmerman

1. INTRODUCTION

1.1 DESCRIPTION

For my research project, I will research and implement a heat map. The implementation will involve using hashRF to create a distance matrix. A distance matrix assigns a value based on how different two trees are from each other. A higher number means there are more differences. I will upload the distance matrix and color coordinate it based on the values in the matrix. I will also make the heat map interactive so that trees can be moved around. The idea is that by using colors and manipulation patterns that were previously undetectable might become more obvious (see section 1.3 for more details).

1.2 PURPOSE

I propose to research, implement, and produce a formal report on the visualization of a heat map to fulfill the requirement of the Distributed Research Experience for Undergraduates (DREU) and the Texas A&M Summer Undergraduates Research for Engineers (SURE), Summer 2009.

1.3 PROJECT GOALS AND IMPLICATIONS:

The heat map produces a colored representation of the distances between given trees. A higher number in a distance matrix would indicate more differences between trees and therefore, a darker color on the heat map. The distance matrix lists the trees arbitrarily. So my program is going to allow the user to move the columns and rows to see if a pattern emerges. It will also allow people to zoom into specific areas of the map. Once these features have been implemented, I will add additional features that I think might be useful.

1.4 PERSONAL GOALS

I am interested in discovering a different area of Bioinformatics. Previously most of my experience in Bioinformatics has dealt with peak alignment issues. A goal is to discover if I am more interested in genetics or the species part of Bioinformatics.

I am also interested in completing a large project. Often during the school year, I only have the opportunity to implement smaller projects or a small part of a larger project due to time limitations. I am looking forward to completing this project.

Finally, I hope that this research project will give me a better understanding of what graduate school will be like and whether or not graduate school is something that I am interested in pursuing.

2. PROGRAM OUTLINE

2.1 APPROACH

Step	Task	Hours
1.	Outline how the program will be implemented.	8
2.	Determine any problematic areas.	8
3.	Adjust original outline to account for problem areas.	12
4.	Implementing the hashRF to Processing Code	16
5.	Implement the display move and zoom features.	50
6.	Test the Program.	-
7.	Debug the program.	-
8.	Repeat steps 6-7 until program runs without issues.	30
9.	Add additional Features	40
10.	Repeat 6-7 until program runs without issues.	20
11.	Complete any needed documentation for the program that was not created during the course of writing the program.	16
12.	Complete poster and Presentation	8
13.	Complete report	16
14.	Complete Website	6
	Total Hours:	230

2.2 APPROXIMATION OF WEEKLY SCHEDULE

Week End

Steps Completed*

End of Week 5 Steps 1 – 4 Completed

End of Week 6 Steps 5 Completed with some testing & debugging

End of Week 7 Step 6-8 & 12

End of Week 8 Steps 9 Completed with some testing & debugging

End of Week 9 Steps 11-12 Completed

End of Week 10 Steps 13-14 Completed

*For steps please refer to section 2.1 Approach

The poster was moved up to week 7 in order to participate in SURE activities.

2.2 WORK SCHEDULE

My work schedule is as follows:

Monday – Friday 8:30 – 5:00 with half hour lunch.

That time will be spent in my office.

This will be adjusted slightly the weeks ending July 25 and August 1 due to being out of town for my sisters wedding. I will either work longer days the days I am here or work from home to ensure that the 40 hour criteria is met.

3. RESOURCES

3.1 METHODS AND MATERIALS

The university has provided everything that I will need for this project. The programs that I intend to use are open source and just need to be downloaded. The printing in Lab HRBB 210 should be adequate for the documentation and reports needed.

4. DELIVERABLES

4.1 ITEMS AND DATES

The following Items will be delivered on the dates specified below:

1. Research Abstract July 30, 2009 by Noon
2. Research Poster August 3, 2009 by 8:00am
3. Poster Presentation August 4, 2009 at 4:00pm
4. Poster Presentation August 9, 2009 at 9:00am
5. Final Website August 19, 2009 by midnight
6. Research Report August 19, 2009 by midnight

5. CONCLUSION

5.1 SUMMARY

DREU and SURE require research to be completed and documented in the form of both a website and a final report. SURE also requires that a poster be created and presented. The create of a phylogenetic heat map viewing program will require me to do research and implement something that has not been done before. Off of the research and implementation that I do, I will be able to fulfill the requirements of DREU and SURE, Summer 2009.

5.2 CONTACT

For more information, please e-mail me at timmerman.16@tamu.edu, or phone 937-726-6717.

6. FACULTY ADDITIONS FROM TIFFANI WILLIAMS

6.1 PROJECT GOALS

6.2 STUDENT MENTEE GOALS

My expectations for Katie include: ... (iii) writing up the results of her work for scientific publication, and most importantly, (iv) having fun.