

VISUALIZING DIVERSITY AMONG A SET OF OBJECTS

Final Report
Summer 2009

Intern

Nicholas Hubbert
Computer Science, Grambling State University

Mentor

Professor Ron Metoyer
Electrical Engineering & Computer Science, Oregon State University

Grad Student

Tuan Pham
Electrical Engineering & Computer Science, Oregon State University

Author

Nicholas Hubbert

Sponsored by
Distributed Research Experiences for Undergraduates
(DREU)

TABLE OF CONTENTS

<u>ACKNOWLEDGMENTS.....</u>	<u>IV</u>
<u>INTRODUCTION.....</u>	<u>V</u>
<u>DISCUSSION OF PROJECT.....</u>	<u>6</u>
<u>FRONTEND.....</u>	<u>8</u>
<u>BACK END.....</u>	<u>11</u>
<u>Conclusion</u>	<u>15</u>
<u>References.....</u>	<u>17</u>
<u>BACK END ERROR: REFERENCE SOURCE NOT FOUND</u>	
<u>Add a User page Error: Reference source not found</u>	
<u>Add a Question page Error: Reference source not found</u>	
<u>CONCLUSION ERROR: REFERENCE SOURCE NOT FOUND</u>	
<u>REFERENCES ERROR: REFERENCE SOURCE NOT FOUND</u>	

LIST OF FIGURES

Fig. 1 Parallel Coordinate Visualization.....	2
Fig. 2 Barycentric Visualization.....	2
Fig. 3 Log-in page.....	4
Fig. 4 Background Info Page.....	5
Fig. 5 Backend page.....	6
Fig. 6 Add a User.....	7
Fig. 7 Add a Question.....	9

ACKNOWLEDGMENTS

I would like to thank DREU for giving me a chance to have such a nice and challenging experience that will help me pursue my scholastic journey. I would also like to give acknowledgement to Professor Metoyer and all of the grad students for not thinking twice to lend out a helping hand and inviting me to contribute on such a great project.

INTRODUCTION

My name is Nicholas Hubbert and I am a junior currently attending Grambling State University majoring in Computer Science. I just completed summer research at Oregon State University under the Distributed Research Experiences for Undergraduates (DREU) program. My learning experience mainly focused on displaying data dynamically and transferring data to and from a database efficiently. The experience exposed me to the rigors of meeting deadlines, communicating within the team, and working with team members to meet an objective.

DISCUSSION OF PROJECT

Have you ever wanted to find out the depth and the diversity over a specific set of objects? For example, consider the problem of determining the amount of diversity in an incoming freshman class of a university, or the roster of a fantasy football team, or the makeup of a stock portfolio. College admissions may find it useful to know, for example, that the incoming class contains 200 males compared to only 20 females which may not be the desired distribution even though both genders are represented in the population and all other properties are possibly well distributed (i.e. GPA, income level, etc.)

In most cases, measuring the overall diversity amongst a set of objects can be examined through a specific set of attributes. For example, one may want to find the GPA, gender, ACT, etc. of a group of incoming freshman. Unfortunately, as the number of attributes and objects increase, analyzing the diversity using text or table representations may become very time consuming and complex. Visually encoding the data can potentially solve this problem. The technique created here at OSU uses a variant of the parallel coordinate's visualization method for multi-dimensional data. This method uses semi-transparent rectangles rather than the traditional poly-lines to represent the objects under consideration. Figure 1 shows a visualization of objects considering 7 attributes (GPA, HIGHEST ACT, ETHNICITY, GENDER, RESIDENCY, INCOME, and FIRST GENERATION) of incoming freshman. As the number of students that fall into a particular value of an attribute increases, the opacity of the rectangle increases. For example, the male bucket of the gender attribute is more opaque than the female bucket because there are more males.

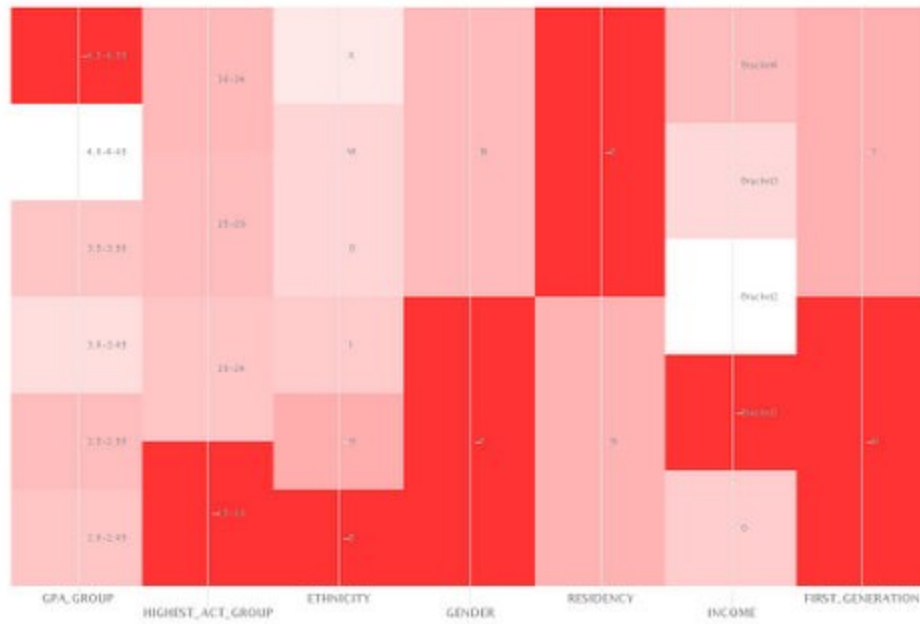


Fig.1

There have been past visualizations techniques such as Fig. 2



The OSU technique will be compared to the barycentric visualization for effectiveness in analyzing the diversity through user studies. My contribution to the project was to help in the design of the user study and to develop an infrastructure for administering the study through a web interface.

Program Goals:

- Take in data efficiently and dynamically
- Display the data taken in
- Allow the user to make choices (i.e. – clicking buttons, submitting answers)
- Store the required data into a database

The infrastructure for the study can be divided into a backend for the study designers to create the study and a front end for the study participants to take the study.

FRONTEND

The front end consisted primarily of the log-in page and a page for participants to input background information that would then be stored in a database.

Log-in page

Creating the log-in page was not that difficult. My mentor wanted the page to only take in a specific User Code and display an error if the code did not exist in the database. This ensures that when a participant sits down to take the study, they have been given an appropriate ID that has been setup beforehand by the study designer. I created this log-in page with two identical pages where one shows an error above the text area where you input the code if it is incorrect.. For example, if the user enters a User Code that does not exist in the database the user will be redirected to the error page that says “The User Code do not exist” and will be asked to enter another User Code.

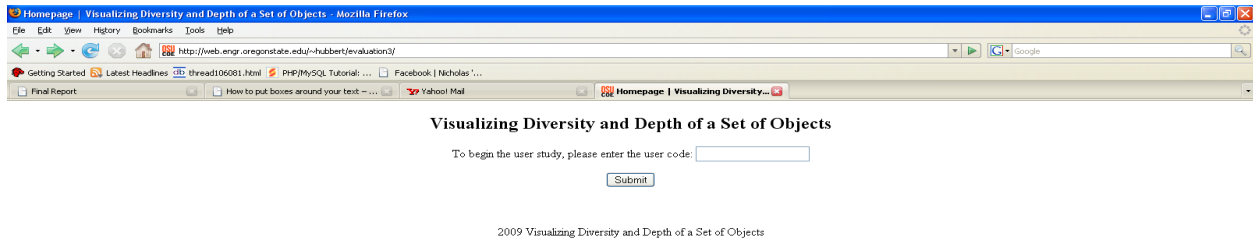


Fig. 3

Background Information Page

In addition to the front end, I was responsible for the background information page. This page is an HTML page that asks the user for their Age, Ethnicity, Major, etc. The information is stored in a database and later output to a background summary page which summarizes every user who took the survey in a simple table. The background summary is useful if the administrator wants to quickly gather the number of males or females, ages, majors, etc.

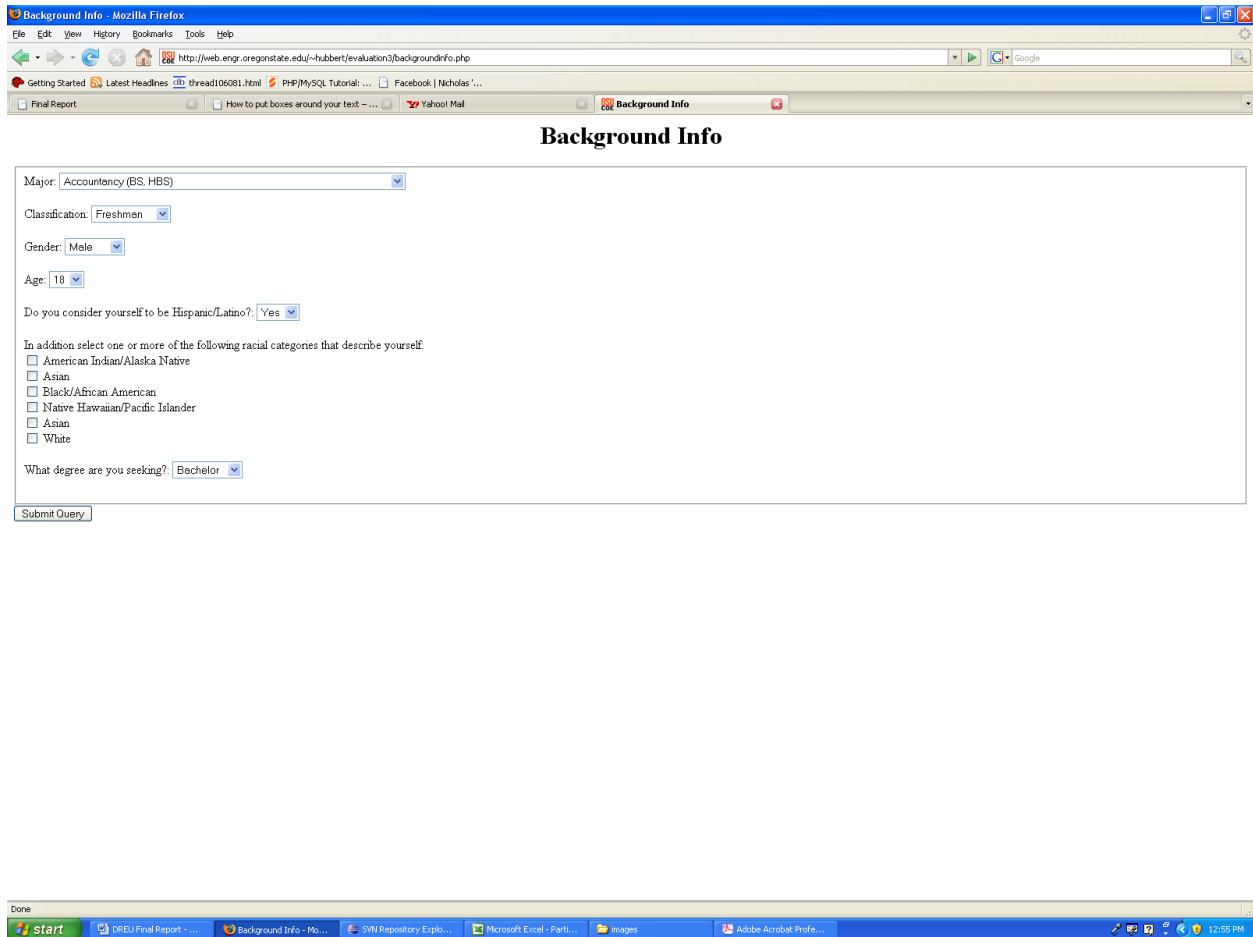
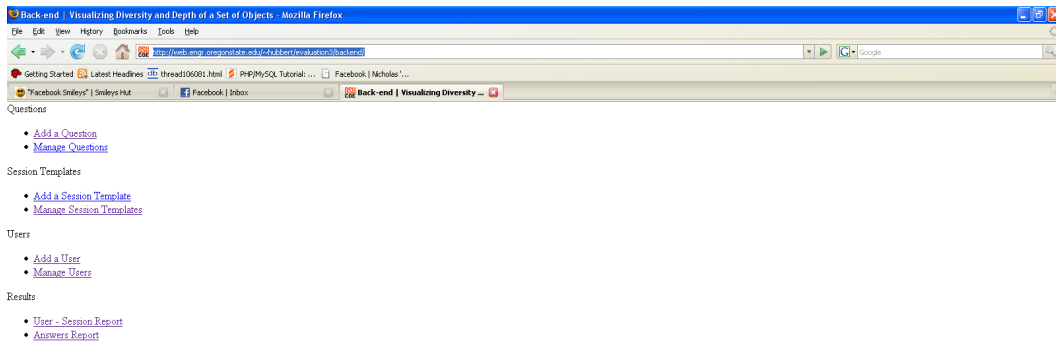


Fig. 4

BACK END

The back end of the program is intended for the study administrator. The page consists of six options, two of which we worked on during the summer: Adding a User, Adding a Question.

These are the most common options that the study administrator would need...



To begin, please select a menu item on the left.

2009 Visualizing Diversity and Depth of a Set of Objects



Fig. 5

Add a User page

I was primarily responsible for the Add a User option. The page was designed to give the user the ability to add a user code and assign a session template I.D without manually accessing the database. The difficult part was creating the query that would assign the Session Template I.D to the user once it is created. A session template consisted of a specific order of questions defined by the study administrator in the back end. We currently have 3 session templates. My intern collaborator, Maduka, who is also a DREU participant, helped me design a query to populate the question table with the User I.D and the questions that corresponded with the Question Template I.D.

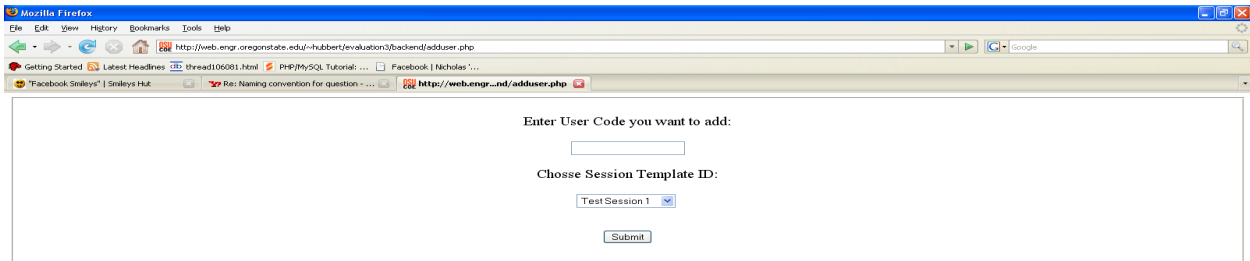


Fig. 6

Add a Question page

The page for adding a question was mostly created by my collaborator, Maduka Ogba (Figure 7). This page gave the administrator the ability to add questions and all the appropriate parameters for the question such as the required images and the question type, (multiple choice, scale 1-10, short answer, etc). The naming convention Tuan, grad student, created for creating names was:

I was mostly responsible for the database administration component of the questions. I named all of the questions using an agreed upon naming convention and entered all of the images into the database using the Add a Question page.

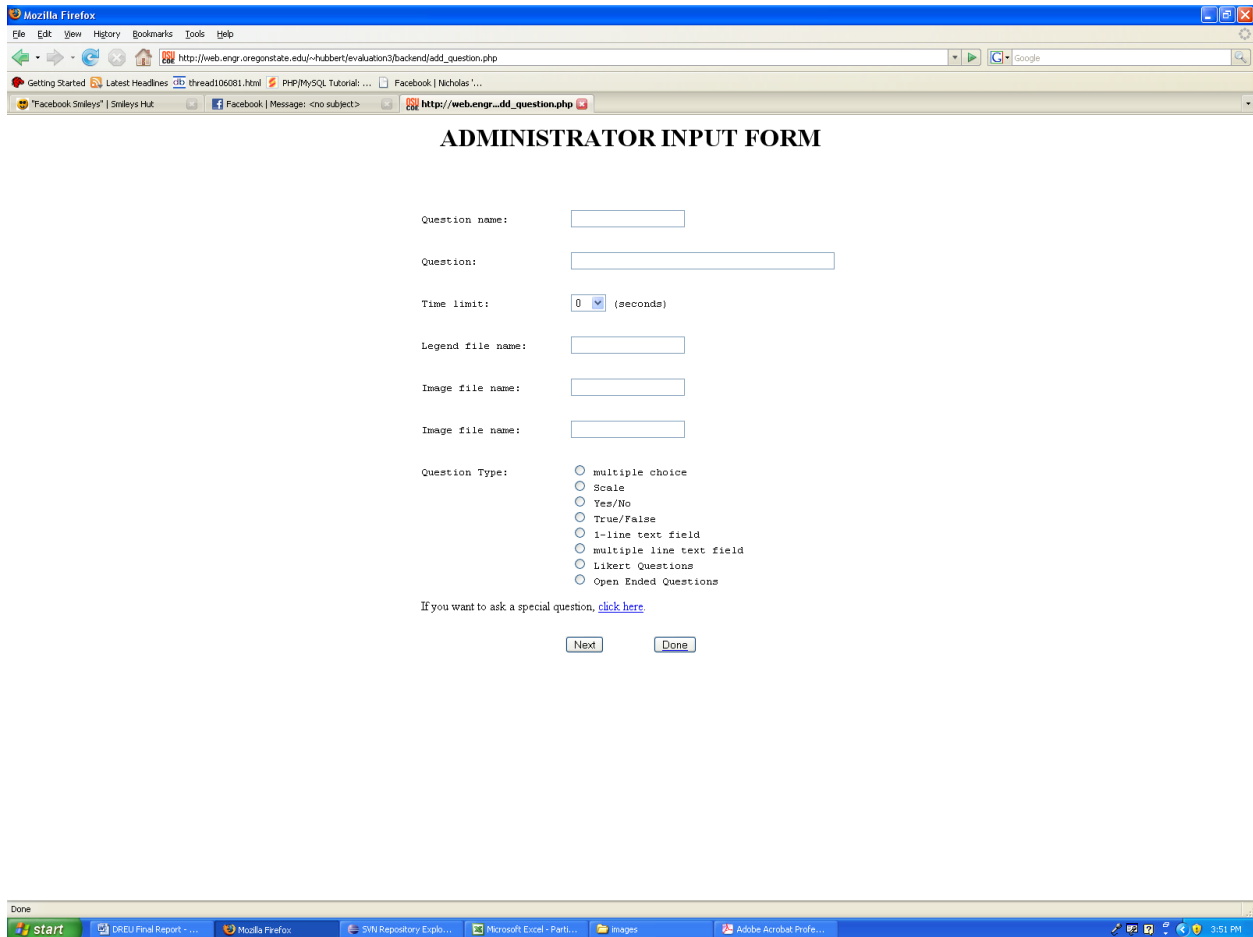


Fig. 7

CONCLUSION

We finished most of the tasks necessary to produce a system capable of running a test pilot. Unfortunately there were a few tasks that we did not have time to finish. Overall, considering that this was my first opportunity to get hands-on experience with research, I believe I did a great job in managing data in a database, displaying data efficiently through a web page, and linking my work with that of the team. We all used a common open source program called Eclipse to create the PHP source code, which was a valuable learning experience for me. I

found the experience with databases rewarding and hope to continue in working with databases and eventually specialize in database management/administration.

REFERENCES

Books

Ulman, Larry 2005. PHP and MySQL for Dynamic Web Sites. California

Reports

J. Pearlman, P. Rheingans, and M. des Jardins. Visualizing diversity and depth over a set of objects. IEEE CG&A, 27(5), 2007

T. Pham, R. Hess, C. Ju, R. Metoyer, J. Gilbert. What Does Diversity Look Like? Oregon State University, 2009

Online Sources

W3Schools Online Web Tutorial, 2009, <http://www.w3schools.com>