Designing Undergraduate Research Through Computer Science Teaching

Yousuo Joseph Zou

Abstract **- How to cultivate undergraduate students' research capability and prepare them for their future graduate study is one of the major tasks of our university's mission. This experimental research was conducted by the author in his teaching practice of computer science in 2004 and 2006 at University of Guam. This paper introduces the author's experiences and lessons in planning, designing and guiding undergraduate students in their research practice and research findings through cooperative learning environment, teamwork and research projects.**

Index Terms – Computer science teaching, planning, design, undergraduate research

I. INTRODUCTION

Undergraduate research is important since it prepares both students for their future employment and their further study in graduate school [1]. In the United States, most of universities and colleges offer their undergraduate students opportunities to gain some sort of research experience [2], just like the Harvard University which established a research program for their undergraduate students, PRISE (Program for Research In Science and Engineering) [3]. Some of these universities have a formal process for placing undergraduate students in research labs or helping their students to find their research professors to do some research rehearsal through internship programs. Through these activities, students have opportunities to chat or interact with graduate students, post-doctoral researchers, faculty members or experienced research experts [4], and they learn research techniques and problem-solving skills as well as scientific culture and career by attending regional and international scientific conferences and present their research findings or publish their research papers in academic journals [5].

According to research findings on undergraduate research, scientific research career training can be classified into five levels: lower division, upper division, master's, doctoral and post-doctoral levels [1]. Therefore, undergraduate research can be planned and designed for the junior or even freshman students in colleges. However, how to plan and design undergraduate research in a specific academic area, such as Computer Science, are still a challenge to university professors. Greening and Kay [6] at University of Sydney published their undergraduate research experience in computer science education, Polack-Wahl and Squire [7] shared their experience on how to overcome obstacles to undergraduate research in computer science at a small institution; Barr and Fekete, et al [8] introduced their research on problem-based learning in undergraduate computer science education; Way [9] used a virtual laboratory model to encourage undergraduate researches, and, Teller and Gates [10] employed research group model to stimulate undergraduate research in computer science.

In this paper, the author wants to share his own experience and lessons in planning, design, and implementation of undergraduate researches in computer science at University of Guam.

II. PLANNING THE UNDERGRADUATE RESEARCH

When you are planning your undergraduate research, first, you need to follow the missions of your university, your college, and your department. Second, you should also consider your students' knowledge background to make sure the planned research will be successful under your guidance. Third, you need to prepare as early as possible the laboratory and the research tools that are required in your planned

Manuscript received December 12, 2009.

Dr. Yousuo Joseph Zou with University of Guam, P. O. Box 5002, UOG Station, Mangilao, GU 96923, USA. (e-mail: yjzou@uguam.uog.edu)

research projects. Actually we started the preparation of computer equipment and computer labs in Autumn 2003 and successfully established 2 computer labs at University of Guam: a Software Development Lab and a Computer Networks Lab in Spring 2005, that made our undergraduate research projects possible starting from the Autumn semester 2005. Fourth, you need to collect real world data from local or regional industrial, business or research institutions. These data are to be processed and analyzed by computers and algorithms. Fifth, you need to plan and prepare the right methodology or algorithm for computers to analyze your data. The quality level of the algorithm you plan to use will directly relate to the acceptance or publication of your research findings by academic conference proceedings, academic books or journals. If you want your students' research findings to be published by world-level journals, your data processing algorithm should also be world-class level. The last thing in planning and preparing your student research is the presentation or publication of the student research findings. Although this is the task after your student research, you need to plan it in advance. Based on the above planning, we worked on our preparation starting from August 2003 and completed the following preparations in Spring 2005.

1. The research findings were expected and planned to be presented or published at international level (above local and regional level).
2. The well installed computers and the Software Development Lab located in the 2nd floor of the School of Nursing and Health Science Building were ready to be used. The lab housed 21 new computers with Windows XP operating systems, Microsoft Office 2003 and successful Internet connection and e-mail service. Each computer in the lab was equipped with software development kit (SDK) and other software execution tools such as MatLab, etc. The SDK platform was able to support several programming languages: C++, Java, Assembly, HTML, FORTRAN, COBOL and Quick-Basic, so that students can select their best programming language to develop their software.

3. The experimental data for the research were all from lab or field investigations in Guam or in the Western Pacific region. Considering the most important industry in Guam is tourism, we cooperated with the Guam Visitor's Bureau of Government of Guam and the International Tourism Research Center at University of Guam and collected the arrival tourist data from 1990 to 2004. This is the 1st data set we used in research. The 2nd dataset is the climate data bank of 1999 – 2004 obtained through collaboration with the Meteorological Observation Stations located in Yigo, in northern Guam Island, and in Inarajan, in southern Guam Island. These two Weather Stations are jointly operated by the University of Guam's College of Natural and Applied Sciences and by the U.S. Department of Agriculture's Experiment Station in Guam. The meteorological data monitored and recorded includes wind speed and wind direction, humidity, rainfall, temperature and solar radiation energy in five-minute averages, hourly-averages, and daily-average values. The 3rd dataset is the top 10 different agricultural products imported into Guam from different foreign countries during 1993 – 2002. The 4th dataset is the workforce survey of personnel and skill demand from 63 business companies/organizations in Guam. The 5th dataset is the field measurement data of mineral element's concentration and its geographic distribution in Guam (different mineral element concentration values in different geographic places in Guam). The 5th dataset were obtained through collaborative research with Saint John's High School's student research team located in Upper Tumon, Guam, USA. These world unique data obtained through field-investigation or lab-experiment already indicated the originality of our designed student researches.

4. After obtaining the high-quality original data, next is the program algorithms to analyze the data; otherwise, the research findings will not come out. Since our research expectation is toward the international level, the mentor of the students' research has carefully designed the data analysis and processing methodologies, or unique computer program algorithms for the undergraduate research projects. The 1st algorithm is digital data file processing, analysis and visualization. The 2nd algorithm is computer modeling and simulation based on the properties of these datasets and the theory of curve-fitting techniques. The modeling and simulation are divided into linear mathematical simulation and non-linear mathematical simulation by solving multi-variable linear and nonlinear algebraic and differential equations using computer software.

5. The students selected into these research projects must have the required knowledge or be able to learn the required knowledge in the designed data processing algorithms and software development. The required mathematics for the designed research are linear algebra, non-linear algebra, calculus, differential equations, basic

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

statistics and probability theory, partial differential equations, numeric analysis. The required computer science knowledge includes traditional and object-oriented programming, operating systems, data structure and algorithm analysis, computer graphics, data file and database management, data communication and computer networks, and introduction to software engineering. Considering the undergraduate computer science curriculum at University of Guam, the selected students for the designed researches should be the senior computer science majors and a few qualified juniors.

III. DESIGN OF THE STUDENT RESEARCH

Based on the above planning and preparation, we were able to design the undergraduate student research in the summer of 2005, and then decided to implement the design for student research in the Autumn semester, 2005 when the author of this paper started to teach a senior computer science course (CS 431 Advanced Topics: Classical and Object-Oriented Software Engineering) at University of Guam. The designed research heavily relied on making use of the above prepared computers and Software Development Lab, software development tools, datasets, and algorithms, with selected senior and junior computer science students.

In the Autumn Semester, 2005, the author approved 17 senior students (15 males and 2 females) majoring in computer science and computer information systems into his CS 431 class. Based on other professors' teaching experience [5, 10] and his personal teaching experience, the cooperative learning environment in the format of research and learning groups was found an effective teaching strategy. Hence, he divided the 17 students into 4 research groups with 4 to 5 students in each group. In discussion with each group, he assigned each a different research project and different data analysis algorithms as follows:

Group 1: develop software to analyze climate dataset, trying to find the quantitative relationship (mathematical formula) of humidity variation with time (daily, monthly and yearly variation) in Guam, trying to find quantitative relations between Guam humidity and temperature using the observed climate dataset.

Group 2: develop software to analyze Guam tourists arrival dataset, trying to find the quantitative (or mathematical) relationship of arrived tourists' number variation with time (monthly and yearly), and also linking to the change of the economic income with time (yearly change) of tourists' home countries, and explaining the reason of the these variations, trying to predict the future trend of Guam's tourism industry.

Group 3: develop software to analyze the climate dataset, trying to find quantitative (mathematical) relationship between solar radiation strength variation with time (daily, monthly and yearly variation). Meanwhile by using statistical methods, try to find the quantitative relationship between solar radiation variation with relative humidity variation in Guam.

Group 4: develop software to analyze the dataset of Guam's top 10 imported agricultural products, trying to find quantitative (mathematical) relationship between arrived tourists' number variation with the consumption of food and other agricultural products yearly that need to be imported (shipped) into Guam, and predicting the future trend of the demand of food and other agricultural products in Guam.

Except for the different dataset and handouts for the programming algorithms and related mathematics given to each group, the above 17 students in 4 groups all used the same textbook, Object-Oriented and Classical Software Engineering, 6th edition by Stephen R. Schach, McGram-Hill publisher @ 2005, ISBN 0-07-286551-2.

In order to use group collaboration to help students develop their research skills and problem-solving capability while also stimulating individual's learning enthusiasm, the class instructor (and also the research mentor) employed the following grading policy: Each student's research project grade consists of two parts: 50% from group-averaged points and 50% from individual's work submitted. Each individual student was required to submit his/her own progress at the following milestone phase of software development cycle: requirement analysis phase, specification phase, design phase, implementation phase, testing phase, documentation and delivery phase. The group-averaged points are the sum of each member's earned points to be divided by the total number of group members. The instructor let each group elect their own group leader and encouraged them to help each other within the group; otherwise, the failure of any member of the group would influence the group-averaged points. This grading policy encourages individual members also work hard under the assistance of the group members. A lazy member will fail by himself/herself since he/she was required to submit his/her own work.

In order to stimulate students' learning enthusiasm by group competition, the instructor (or research mentor)

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

announced the following bonus points to award the leading group at the beginning of the semester: after the research project was finished, the group with the highest group-averaged points, each member would receive 5 bonus points; the group with 2nd place, each member would receive 3 bonus points; and 3rd place, each member would receive 1 bonus point; and last place, each would receive 0 point.

The grading policy for the whole class was: research project 45%, attendance and class performance 10%, midterm exam 20%, and final exam 25%. Students without proper programming skills or necessary mathematical knowledge usually failed in these two exams.

The above 1st experiment for student research was successful. All the 4 groups completed their research and achieved satisfactory research results. The details of the research findings are described in the next section.

Based on the success of the 1st experiment in Autumn 2005, the 2nd experiment for the student researches were also designed and implemented in the Autumn semester, 2006, in which 16 computer seniors were approved into the CS 431 student research class. This time, the author of this paper divided them into 6 research groups. Each group only had 2 or 3 members. The instructor/research mentor employed the same textbook, same handouts and same grading policy but assigned a new research task to each of the 6 groups by developing related software to analyze their different dataset and harvest their different scientific findings.

Group A project: use the Guam tourist arrival dataset and the dataset of Guam's top 10 imported agricultural products to study the influence of tourism industry on the demand of agricultural products in Guam.

Group B project: use Guam tourist arrival dataset in 1994 – 2002 to analyze the characteristics of Guam tourism industry.

Group C project: use Guam climate dataset to find quantitative relationship between humidity and rainfall variation.

Group D project: using the mineral elements' dataset to study the distribution of mineral elements in Guam and their relation to the cancer diseases of Guam patients.

Group E project: use Guam climate dataset to analyze the temperature variation with time and predict Guam's temperature change.

Group F project: use Guam workforce dataset to analyze the Guam employers' demand of specific professionals with specific skills, and predict Guam's future job market.

Basically speaking, following the above-described design, the 6 undergraduate students' research projects were also successful. The details of the research results will be given in the following section.

IV. RESULTS OF THE EXPERIMENTAL STUDENT RESEARCH

After implementation of the above design in the Autumn semesters in 2005 and 2006, the 10 designed student research projects were all successful. The author of this paper then made his effort to help the student researchers write their research papers on the above-described 10 student research projects implemented in Autumn 2005 and Autumn 2006 and sent these research findings to local and international academic conferences. Among the above 10 students' research projects, 6 papers had been accepted and presented at an Annual Research Conference of University of Guam, March 12 – 13, 2007. The following briefly describes the accepted student research papers.

Student Paper 1: A computer software that analyzed the characteristics of Guam's tourism
Industry during 1994 – 2002, student authors: L. A. Manalang, V. Abiera; mentors: Y. J. Zou and J. C. Salas.

Student Paper 2: The influence of the tourism industry on the demand of agricultural products in Guam – A quantitative study of using newly developed software, student authors: C. Farnum, A. Blas, and M. Helgenberger; mentors: Y. J. Zou and B. Barber.

Student Paper 3: Analysis of Guam's workforce data based on self-developed software, Student authors: R. A. Ulangco and A. Tesalona; Mentors: Y. J. Zou and A. B. Enriquez

Student Paper 4: Prediction of Guam's temperature by self-developed software, Student authors: A. Umayam, D. Sablan, and R. Itugot; Mentor: Y. J. Zou.

Student Paper 5: Data analysis software and the humidity – rainfall relationship of Guam,
Student authors: P. V. Rabino, G. Ocampo, and R. Ada; Mentors: Y. J. Zou and P. Singh

Student Paper 6: A study of the distribution of mineral elements in Guam via software, Student authors: E. Del Rosario and B. Biacan; Mentors: Y. J. Zou and C. Evola

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

The above 6 student research papers were also accepted and presented at the International Conference on Business, Economics, and Information Technology, March 5 – 6, 2007, Guam, USA. Please refer to the conference web-site: www.altoona.psu.edu/icbeit/hanoi/docs/GuamReport2007.pdf

We are still working with these graduated students to submit their papers and other students' papers to be published in academic journals.

V. CONCLUSION

The reported work on undergraduate research during 2004 - 2006 at University of Guam made the author gain some successful experience but also some lessons. For example, the above designed research topics with algorithms for data analysis may not be interested in by every student. Some students wanted to select their own research topics but we could not satisfy their demands partly due to the lack of hardware or software equipment or research funds, partly due to lack of planning and preparation time. Hence we need to plan more and prepare more research topics to meet students' needs. In some regional universities, efforts on undergraduate research come across resistance from administrators or other faculty members and even students since they have not realized the importance of student research.

The author hopes his experience and lessons can be shared by other mentors of undergraduate research.

REFERENCES

[1]. C. Gonzalez, Undergraduate Research Graduate Mentoring, and the University's Mission, Science, 2001, vol. 293, 1624 – 1626.

[2]. S. Goodland, Research Opportunities for Undergraduates, Studies in Higher Education, 1998, vol. 23 (3), pp 349 -355.

[3]. A. D'Gama, Got your Science On: With the Harvard Program for Research in Science and Engineering (PRISE), The Harvard University Research Journal, 2009, vol. 2(2), pp. 11 – 12.

[4]. C. M. Kardash, Evaluation of Undergraduate Research Experience: Perceptions of Undergraduate Interns and Their Faculty Mentors, Journal of Educational Psychology, 2000, 92 (1).

[5]. M. Fischer, Chess 2.0: A Web Experiment in Collaboration, Stanford Undergraduate Research Journal, 2007, vol.6, pp 43 – 45.

[6]. T. Greening and J. Kay, Undergraduate Research Experience in Computer Science Education, Proceedings of the 7th Annual Conference on Innovation and Technology in Computer Science Education, Session of Learning Experience, 2002, pp 151 – 155, ACM Publications, New York, NY, USA, ISBN:1-58113-499-1.

[7]. J. A. Polack Wahl and P. N. Squire, Overcoming obstacles to undergraduate research at a small institution, Journal of Computer Sciences in Colleges, 2003, vol.18 (5), pp 128 -135.

[8]. M. Barr, A. Fekete, T. Greening, O. Hollands, J. Kay and J. H. Kingston, Problem-based learning for foundation computer science courses, Computer Science Education, 2000, vol. 10(2), pp 1 – 20.

[9]. T. P. Way, A Virtual Laboratory Model for Encouraging Undergraduate Research, ACM SIGCSE Bulletin, 2006, vol.38 (1).

[10]. P. J. Teller and A. Q. Gates, Using the Affinity Research Group model to involve undergraduate students in computer science research, J. Engineering Education, 2001.

ISBN: 978-988-17012-8-2
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)