Experiences Teaching a Graduate Research Methods Course

Clifford A. Shaffer
Department of Computer Science
Virginia Tech
Blacksburg, Virginia 24061 USA
shaffer@vt.edu

Abstract

Teaching research methods to computer science graduate students presents a number of significant challenges. Of particular concern is the extreme breadth of material that must be covered and the diversity of the students taking the course. This paper describes the implementation of the graduate research methods course at Virginia tech. The primary goal of this course is to improve the ability of research-track students to perform graduate research and more generally to prepare for their professional career (not necessarily in research). A notable feature of our implementation is the use of authentic assignments whenever possible.

Keywords: Research methods, technical writing, experimental design, statistics, professionalism

1. Introduction

A fundamental problem faced by Computer Science graduate programs both at Virginia Tech and at many other universities is the inadequate preparation of most graduate students for graduate-level research. Most students are deficient in one or more of the following areas: technical writing skills, written and oral presentation skills, basic statistics, and a general understanding of how the research ``game'' is played. While some students are knowledgeable in one or more of these areas, almost none of our students have any training in experimental design. This is a significant deficiency for our program, since much of the research focus of our department is oriented toward either experimental computational science or HCI, both of which make heavy use of such material. All of these topics are likely to be important for students while doing graduate-level research, and even more so once they graduate and enter the profession, regardless of whether their career is specifically oriented toward research or not.

The most common ``solution'' to these deficiencies is to leave students to absorb this information during the natural process of their graduate student career, with any remaining gaps made up through on-the-job training by the research advisor. Some aspects of professionalism, writing, and presentation skills are best learned as part of the mentoring that characterizes the graduate-level research process. However, both the research programs and the students would be better off if greater proficiency in these topics were achieved prior to embarking on the first research project, or at least taught in a more efficient manner than one-on-one interaction with a research advisor.

To better prepare our students for their graduate student career and their future professional career, we offer a course entitled Research Methods in Computer Science. Implementing this course presents a number of unique challenges. The two most obvious ones are that the body of material to be covered is extremely broad, and that the students taking the course have a great variation in relevant background. Some topics we cover are themselves often presented as full semester courses (or even two semester sequences). On the surface, the idea of teaching anything related to technical writing at the graduate level seems doomed, considering that the students should typically have had sixteen years of prior instruction in writing.

Our Research Methods course has the unique distinction of being the most controversial course in our curriculum among the faculty. One (growing) vocal minority is strongly supportive of the course since it is perceived as addressing real and pressing problems, and this group feels that any additional exposure to the relevant topics must result in some improvement in student abilities. Another equally vocal (but shrinking) minority objects to the course because it is so broad that it cannot have a practical impact. Further, they view the material as remedial and so oppose the notion that it deserves graduate-level credit in Computer Science. Regardless of this controversy, the Research Methods course receives consistently strong enrollments and it appears to have become a central pillar of our graduate curriculum.

2. Goals and Challenges

The goal of our Research Methods course is first to improve the ability of our graduate students to do graduate-level research and second to prepare them for a professional career in computing. Both of these goals have much in common, in that graduate-level research and a professional career will both be enhanced by improvements in technical writing skills, presentation skills, basic statistics competence, and some proficiency in experimental design. Specific skills such as proposal writing, academic paper presentation, and a general understanding of the academic profession are useful to all computing professionals and to researchers in particular.

Implementations for such a course must deal with two key challenges. The first is that the students populating the course are extremely diverse in ways that directly impact the course. Perhaps the most obvious is the mix of native and non-native English speakers, which most directly affects students' proficiency in technical writing. Even students whose primary language is English will often be weak at technical writing, and students for whom English is a second language tend to be even weaker. Beyond differences in English skills, the next most obvious difference is proficiency and affinity for mathematics in general, and statistics in particular. A significant percentage of our course is taken up by technical topics on basic statistics and experimental design. Some students come to the course already proficient in statistics, but many do not. Thus, some students will not learn much from this section of the course (since they already know that material), and others will find it quite difficult because their background is weak. Finally, some students come to the course with substantial experience in ``doing research'' and might already have published papers. Other students have no experience whatsoever with academia beyond sitting through undergraduate classes, and in particular have never been exposed to any significant research endeavor.

Given the long list of topics that we hope to cover, with some of those worthy of their own course or even whole curricula, it is important to decide just what we hope to get across to the students. It is impossible to ``teach'' technical writing in any depth within the time limit of one-half or one third of a semester course. Likewise, we cannot provide the equivalent of a semester course in statistics, a semester course in research design, and a semester course in professionalism. However, with careful planning, we successfully deliver a course that the students find rewarding and which we believe does have lasting value for the students' education.

3. Content

While I do not partition the course calendar in this way, it is useful to divide the course content into three broad areas: Technical writing and presentation, professionalism, and statistics and experimental design. There is much interleaving of the first two areas, and some subtopics might arguably be a member of either area. Each area takes up about one third of the semester

3.1 Technical writing and presentation

Technical writing and presentation topics are the most problematic. A one-semester course cannot make up for deficiencies from sixteen years of education or make fluent writers out of non-native speakers. For example, it is not practical to teach English grammar in this course. Nonetheless, there is useful information to convey. I use the standard strategy of making students write, and critique it, since the more direct feedback on writing skills a student receives, the better. There are particular ``hot spots'' in writing that we focus on. These include minimizing passive voice, proper use of articles for non-native speakers, and reducing use of words like ``very.'' More generally, I focus on improving clarity, reducing cognitive load on the reader, and being concise. My writing exercises especially attempt to push students toward eliminating use of passive voice, and toward taking deliberate effort to shorten the number of words needed to convey the message. Some of the assignments (described in Section 4) are specifically designed to give practice in revising and shortening text. In all cases, I encourage students who need help with their writing to make use of the Virginia Tech's Writing Center. Occasionally I will encourage (or even require) foreign students with serious deficiencies in English proficiency to take additional courses.

Merging into the professionalism area of the course, I try to instill a philosophy of writing where every document has the purpose of achieving a desired effect on a specific audience (convey information of significance to the audience, convince the audience of something), rather than writing to cover material. I stress the concept that audiences must be hooked at an early stage, and then interest must be maintained. I make it clear that much of academic writing involves meeting a specified format, outline, length, and style.

Another important topic is the process of implementing a paper. This includes the writing and revising cycle, organizing content for a paper, and understanding the role of all the major structural elements of an academic paper. I explain that professional writing demands suitable tools such as document editors that support cross-referencing and that provide suitable bibliography tools. I try to draw as many analogies as I can between writing computer programs and writing technical papers. Not only are there many aspects in common to the management of these two endeavors, but I think that students who are drawn to computer programming will find it easier to tackle their writing tasks if they use their same practiced skill sets. For many students, writing is an unappreciated task, and the realization that many of the skills they apply to a task they find enjoyable

(programming) can also be applied to writing can hopefully make writing more attractive to them.

I spend significant class time on proposal writing, since both academic and industry professionals spend a lot of time in this process. I draw on materials provided by the University on how to write successful (i.e., winning) proposals for my presentations of this topic.

We discuss how to give a presentation, including both producing the materials and delivering the presentation itself. I acquaint students with the university library and its online resources, through presentations by our library staff. Primary online resources include the library catalog, citation indices, abstract services, and journal archives. Finally, I give some advice on how to survive the process of writing a thesis or dissertation.

3.2 Professionalism

The second content area covered by the course is ``professionalism." This includes a wide variety of topics, many of which are closely associated with the material presented on writing and presentation. A common theme running throughout the discussion is evaluating quality: evaluating research, evaluating papers and proposals, and evaluating students (or employees). Judging quality is a major part of academic life.

An important goal of the course is to make explicit what being a research-oriented graduate student, and a research professional, is all about. We seek to motivate students to model being such a person. A key part of recent restructuring for our entire graduate program is to get students involved in the research process early. We tell them to ``start acting like a researcher" and explain what that means. The Research Methods course plays a key role in our program by giving them the knowledge and skills necessary to make this a reality. Ultimately we seek to help them catch a vision for research and what it looks like in Computer Science.

We begin the course with a discussion of why students go to graduate school and how they can be successful as graduate students. All graduate students in the program are encouraged to attend the second lecture of the semester, where I discuss topics such as selecting and dealing with advisors and surviving the PhD qualifier process.

Other topics under the area of professionalism include discussions of general academic life including the various options for academic positions, the life of faculty members, and the tenure system. We spend a couple of weeks on topics related to how to get a paper published (picking a venue, the publication cycle), how papers are reviewed (and how to be a reviewer), and likewise the professionalism aspects to writing proposals and getting them funded, along with reviewing proposals. Finally, we discuss the process of defending a thesis (as opposed to writing the document).

3.3 Statistics and Experimental Design

The third broad area I cover is quite different from the other two, and far more technical: statistics and experimental design. Careful planning is needed to decide what should go into the limited time available for these topics. Each deserves at least a full semester of its own. The theme that I use to convey the material is that of performance evaluation and being able to carry out the type of experimental work that is typical for Computer Scientists. Partly this is influenced by the textbook that I use [2], which stresses developing a methodology for performance evaluation.

The statistics topics focus on concepts related to hypothesis testing. Many branches of Computer Science require students to compare performance or populations to determine their relative qualities. It is quite frequent that computer science researchers ask which algorithm, program, or system performs better in some way, and I try to motivate all statistical techniques by clearly focusing on how they relate to this fundamental question. Specific topics we cover include measures of central tendency, correlation and covariance, indices of dispersion, sampling, hypothesis testing, confidence intervals, t-tests, regression (mainly simple linear regression and some multiple regression), allocation of variation, ANOVA, and cross tabulations (Chi-Square).

While many students come to our program with some exposure to statistics, it is rare that they come with any exposure at all to experimental design techniques. Yet, many of our students are doing formal experimental work in research areas such as HCI and networking. So we include material on formal experimental design (in fact, this topic was a primary motivator for the original creation of the course). Topics include full factorial, fractional factorial, 2^k factorial, and 2-factor full factorial experiment design techniques.

Table 1 shows the weekly schedule that I used when I taught the course in fall 2005. Further implementation details can be found at the course website [5]. I used two textbooks to support this course. Material on technical writing and professionalism came largely from Higham [1], which does a good job of blending these areas together. For material on statistics and experimental design, I relied on Jain [2]. While Jain's content is good, the book suffers from a poor writing style (though we can make the most of this by using it as an illustration to the class about some things not to do in technical writing).

4. Assignments

Knowing the topics covered gives only part of the ``feel" of a course. Coming up with meaningful assignments for this course is unusually challenging. The technical portions of the course (statistics and experimental design) are fairly straightforward in terms of appropriate assignments. In a typical semester I give seven problems (for a total of 17.5% of the total semester grade) related to statistics and

experimental design. The rest of the assignments were oriented toward writing and professionalism in various ways.

An important goal I have for the course assignments is that they be as authentic as possible. Here, ``authentic'' means that an assignment is directly relevant and intrinsically useful to the students' real life and activities, as opposed to an artificial exercise for the purpose of developing proficiency and assigning a grade for the course. In particular, I am hoping that the assignments can help further the students' research activities. Over time, I have consistently strived to improve the assignments with respect to this goal.

Table 1: Weekly course schedule for Fall 2005

Week	*Topics*
1	What researchers do, succeeding as a grad student
2	Tools and techniques for paper writing
3	Technical Engish
4	Revising and publishing papers
5	Publishing and reviewing
6	Using library & online document archives
7	Reviewing and proposal writing
8	Proposal writing
9	Performance evaluation and statistics
10	Statistics
11	Statistics
12	Experimental design
13	Experimental design, presenting quantitative data
14	Giving talks and designing websites

The biggest assignment of the semester (worth 20% of the semester grade) was to write a one-page research proposal. The assignment specification requires that the proposal be for a project that the student intends to do as part of the graduate degree process. In advance of delivering the final assignment, students submit a mini-assignment that is an outline of the final proposal, to insure that they are each on track to completing the final proposal successfully.

Since there is wide variation in progress toward degree within the class, the requirements for the research proposal's contents have to be flexible. The more advanced students already have a set research plan, so producing a proposal was not a difficult task. These are the ones for whom this assignment might be the most directly beneficial toward advancing their research, since they typically had not yet written a proposal for their work. In our department, only PhD students are required as part of the normal degree requirements to write a proposal, and that is usually not done until they are well into the program. Students taking this class have rarely advanced to that stage in their career. For many of the PhD-bound students, this assignment presents an opportunity to start working toward such a proposal. For the Masters students, this assignment presents an opportunity to flesh out and think through their research plans, which to that point likely consist only of an informal agreement with the advisor.

The most problematic students were the ones earliest in their graduate career. A few are even part of our 5-year BS/MS program and technically not even finished with their undergraduate degree. These students often have not reached the stage of selecting an advisor or determining a research area. For these students, the assignment spurs them toward reaching those benchmarks. However, the timing was somewhat premature for them. See Section 5 regarding when this course is best taken by students.

I give a variety of other assignments in addition to the research proposal and the technical problems. The first assignment of the semester is a series of exercises related to writing topics, including detecting and eliminating passive voice, consistent use of parallelism, and reducing wordiness. The original version of these exercises is available on the Internet from Ian Johnston [3]. I also used an assignment from our research librarian that provided practice with various library resources.

As a lead up to the major proposal assignment, I require each student to write a 3-5 page literature review. The first year that I offered the course, I set a specific list of papers to be surveyed. However, I have since allowed the students to survey papers relevant to their own research, with better results. Prior to the literature review assignment, students write an abstract for a paper of their choice (with the constraint that it not have an abstract already), and to write a peer review for a paper of their choice [5].

Another series of assignments started off by requiring the students to write a three page paper on their research plans and aspirations (as opposed to the proposal for a specific project). Students were permitted to include some background on their research experiences, previous degrees, why they came to Virginia Tech, their planned research topic, and their job aspirations after graduating. The specifications for the assignment gave fairly precise formatting instructions (in the spirit of what might be required by a conference proceedings). I marked up the initial version of the papers with typical copy-edit symbols (a topic discussed in class), to give them feedback on their writing style. I then gave another assignment where they copy edited their own papers for improvement, with a particular emphasis on shortening the paper. The third assignment of the series required them to rewrite the paper in half the original allotted space, while attempting to preserve the complete content of the original paper as much as possible. For most students (as would be true for most writers anywhere), preserving the full content in half the space is a relatively easy task!

One other assignment given is to write a one half to one page discussion on potential publication venues for their future research work. A goal of this assignment was

to make students aware of the journals and conferences in their particular research field. Students were required to evaluate the quality or reputation of the various venues, and to discuss their aspirations/expectations for success in publishing at these venues. In one research group this assignment proved so popular that the lab head required all of his students to participate!

The last assignment of the semester requires students to prepare a personal website highlighting their research achievements and documenting their progress as a graduate student. This assignment was originally proposed by our department head, who was concerned with documenting the accomplishments of the graduate students. Ideally, all graduate students in all academic departments would maintain such a website, as an online form of resume.

One important topic is explicitly not addressed by assignments. Ideally, students in a course such as this should be given an opportunity to practice oral presentation skills. However, I decided that devoting class time to the delivery of such presentations was too costly given the competing demands of the course. Thus, there was no oral presentation assignment.

5. Conclusions

More than most graduate-level computer science courses, the research methods course is open to wide interpretation in both presentation style and course content. Students might come to this course with widely differing expectations, as might faculty. Our implementation in particular attempts to cover a broad variety of topics. For these reasons, it would be easy for this course to fail. However, our implementation appears to be successful, as evidenced by its sustained ability to attract students and on feedback that students give for the course. This course has been an important component in restructuring our graduate program to be more focused on research. As a result of our restructuring, in just two years we have come from a program where relatively few MS students did a thesis and only about a quarter of the graduate population was in the PhD program, to where now a significant majority (and growing) of MS students complete a thesis and over a third the students (and growing) are in the PhD program.

One particular problem that must be overcome is the perception by students that this is a ``fluff'' course to be taken for an easy grade. That could arise due to a perception that it is ``non-technical'' and therefore hard to evaluate or that it will not contain substantive and evaluatable assignments. In my own implementation, I attempt to require a level of effort equivalent to other first-year graduate courses. I also attempt to give grades with an average and distribution similar to what I give for my other first-year graduate classes.

The results of the Fall 2005 semester were particularly dramatic. More than one student was prompted by the course assignments to seriously reevaluate their planned research direction and make significant changes for the better. Unfortunately, two or three other students concluded that a research-track degree was not appropriate for them. In each of these particular cases, it was clearly better for them to realized this sooner rather than later.

A significant issue that we struggle with is when to offer the course, and who to offer it to. For the first two years, the course was taught in fall semester. This means that the typical student will take it either in their first semester, or in their third semester of the program. For PhD students, the third semester is a reasonable choice, although earlier might be better. For masters students, the third semester is rather late in the program. For any student, the first semester could be viewed as early, since they usually are not prepared to dive into making the decisions about their research program required by many of the assignments that I use. For this reason, we chose in 2004/5 academic year to offer the course in the Spring semester. The hope was that most research-bound students who enter the program in Fall of a given year would then take it in the following spring (the students' typical second semester). This was hoped to provide the right balance of being late enough that the assignments encourage and support their research progress while not being so late that many of the short-term benefits are lost for starting their research program. However, after trying the Spring semester course offering for one year, we concluded that getting the students started in their research earlier is an overriding consideration. For AY 2005/6 we switched back to a Fall offering, and intend to remain on that calendar.

Acknowledgments

My thanks to Marc Abrams and Cal Ribbens, the previous instructors for CS5014, for many ideas on how to run a graduate-level research methods course. I believe Marc was responsible for the original conception of experimental design as a key component for this course. Each of us has shaped the course in our own ways over the years, but without their efforts to build on, I doubt my version would have succeeded so well as theirs.

References

[1] N.J. Higham. *Handbook of Writing for the Mathematical Sciences*. SIAM, 1998.
[2] R.Jain. The Art of Computer Systems Performance Analysis. Wiley, 1991.
[3] I.C. Johnston. Exercises for review of grammer and punctuation. Available at www.mala.bc.ca/~johnsto/grammar/review.htm, 2003.
[4] J.Palsberg and S.J. Baxter. Teaching reviewing to graduate students. *Communications of the ACM*, 45(12):22-24, Dec. 2002.
[5] C.A. Shaffer. CS5014 Research Methods Course Website. Available at `courses.cs.vt.edu/~cs5014`, 2002.