Working Paper Series
ISSN 1170-487X

Getting research students started: a tale of two courses

by Ian H Witten and Timothy C Bell

Working Paper 92/5
July, 1992

Department of Computer Science
The University of Waikato
Private Bag 3105
Hamilton, New Zealand

Getting research students started: a tale of two courses

Subject area: Research training

Introduction

As graduate programs in Computer Science grow and mature and undergraduate populations stabilize, an increasing proportion of our resources is being devoted to the training of researchers in the field. Many inefficiencies are evident in our graduate programs. These include undesirably long average times to thesis completion, students' poor work habits and general lack of professionalism, and the unnecessary duplication of having supervisors introduce their students individually to the basics of research. Solving these problems requires specifically targeted education to get students started in their graduate research and introduce them to the skills and tools needed to complete it efficiently and effectively.

We have used two different approaches in our respective departments. One is a (half-) credit course on research skills; the other a one-week intensive non-credit "survival course" at the beginning of the year. The advantage of the former is the opportunity to cover material in depth and for students to practice their skills; the latter is much less demanding on students and is easier to fit into an existing graduate program.

The aims of both courses are:

- to provide an introduction to and survey of research being undertaken in the department;
- to develop students' professional skills; and
- to accelerate students' progress in the graduate program.

Because of time pressure, the short course tends to focus on the more practical aspects of the second topic. The next section identifies what one might wish to teach; then we outline the particular versions we have taught, which still only cover a subset of the material.

Professional skills for the computer science researcher

Many research skills—such as knowledge of the literature, what's worth reading and what isn't, a nose for worthwhile problems, good research taste—can only be learned under the direction of a supervisor who is expert in the subject matter. However, our courses are predicated on the belief that researchers need some knowledge and skills that are more general than this.

WHAT IS RESEARCH IN COMPUTER SCIENCE?

Students entering the graduate program tend to be rather mystified by what constitutes scientific research. Although it is hard to give a definitive characterization of what research is and is not, it is useful at least to broach the subject for discussion. The nature of research is probably best conveyed in a concrete setting by examining particular pieces of work. Computer Science is different from other subjects because of its youth and because existing paradigms and lines of enquiry are much more fluid and less firmly established. As Brooks (1978) argues, we work in a tractable medium and build from "pure thought-stuff"; our research is open, unfettered, creative—and, as a result, extremely hard to control. The main message we try to get across is that "writing a program is not (by itself) research."

THE LOCAL RESEARCH ENVIRONMENT

One of our aims is to introduce students to the research being pursued in the department. A balanced survey of the subject as a whole is not required; we merely seek to germinate an awareness of what is going on locally. Students need to be able to identify prospective supervisors, broaden their horizons, be encouraged to interact with each other, and see examples of seminar presentation styles and techniques. This can be usually be achieved by an ordinary program of seminars given by faculty and senior graduate students, supplemented by a brief overview of research in the department given at the time that students are choosing their research topics.

MAKING PRESENTATIONS

Students need experience in giving presentations—for some this class may provide their first experience in front of an audience. Practice at presentation, and feedback from a sympathetic audience, are invaluable, as are class discussions on seminar presentation styles. Students must learn basic presentation techniques and the use of audio-visual

aids: useful sources include Parberry (1988), Mental (1971), Michaelson (1982), and Cleese's (1976) movie.

CRITICAL EVALUATION

Developing skills of critical evaluation is a major aim of graduate study. Students need to be exposed to reviews (such as in *Computing Reviews*), and to learn to read and critically review literature themselves.

THESIS PREPARATION

It is useful to review routine information about what goes into a thesis: the formal sections into which it is divided; the abstract, table of contents, list of figures; the role of appendices; the difference between a reference list and a bibliography; and so on (e.g. Campbell & Ballou, 1974). The level at which the thesis should be written—that is, the targeted audience—and its style—namely formal, scientific description—can be discussed.

THESIS EXAMINING

Students should be made aware of the process by which their thesis will be examined. They often do not realize that examiners are not compelled to read every word of a thesis in order to form an opinion of its value and quality, and that it may not be read in one sitting, or in order. Reflecting on the realities of the evaluation process helps to highlight the crucial importance of presentation and layout in giving a subjective impression of quality and attention to detail. The special importance of the title, abstract, introduction and conclusion are also worth noting. Thesis orals are a nerve-racking experience for students, and can be made easier and more productive by familiarizing students in advance with the oral situation and discussing the regulations and conventions that govern the conduct of orals.

THE PUBLICATION BUSINESS

Research students need to be exposed to the academic publication business. This includes the relative merits of submitting work to journals, conferences, and workshops, the standing of different journals and conferences and how this might be measured, citation indexes, the "CV value" of a paper, and the role of books and edited collections. It also includes the publication process for journals, conferences, and books: calls for

papers, information for authors, the mechanics of preparing and submitting a paper, refereeing, the role of the editor, likely publication delays, procedures for notification of decision, possible re-submission, acceptance, copy-editing, proof-reading, and index preparation. The refereeing process also deserves attention, involving questions of critical evaluation, professional ethics, and one's involvement in and obligation to the scientific community. Parberry's (1989) and Smith's (1990) articles form an excellent foundation for discussion.

ETHICAL ISSUES

There are numerous ethical and societal issues in the scientific publication business, and in research generally. An interesting, controversial, and occasionally outrageous account of the publication "game" is available (Anonymous, 1987), along with accompanying comments by respected scientific figures; this can be used as a basis for class discussion. Crutcher (1991) gives a sardonic view of how scientists can advance their career with minimum effort. Intellectual honesty, including delicate questions of misrepresentation and even fraud, may be addressed (Crawford & Stucki, 1990).

SCIENTIFIC WRITING

We have found that our students need some education in scientific writing, scholarly style, and proof-reading. They benefit from a review of techniques of formal writing (e.g. use of passive voice) as well as common grammatical errors; resources such as Strunk & White (1979) are invaluable here. The planning of writing and the structure of a paper should also be covered, as well as more technical aspects such as citing references. Useful guides are van Leunen's *Handbook for scholars* (1986) and the standard *Chicago manual of style* (University of Chicago Press, 1982); see also Williams (1990).

OBTAINING INFORMATION

Students must learn to use the various abstracting and indexing publications of our field—the *ACM Guide to the Computing Literature*, *Computer Abstracts*, *Computing Reviews*, *Current Contents*, and the *Science Citation Index*. One possibility is to have a librarian explain the library's computer science collection; another is to organize a scavenger hunt in the library to acquaint students with the cataloging system and the holdings in our discipline. Some interesting exercises are suggested by Habeshaw &

Steeds (1987). It is becoming increasingly important to use the Internet to obtain information and monitor recent developments. For this one needs basic net etiquette skills, use of mail, and a knowledge of resources (such as Archie) for locating information. Even more mundane but nevertheless essential, students need to be informed about other communication facilities—rules for long-distance phone and fax use, and postal mail services.

RESEARCHER'S TOOLS AND TECHNIQUES

Students need to learn how to use a few mundane but essential research tools. These include a comprehensive document preparation system and drawing program, bibliography database, spelling checker, and the like. Depending on the students' interests, other tools, such as statistical analysis programs and symbolic algebra packages, can be introduced. Ideas such as exploratory data analysis in the style of Tukey (1977) and graphical information presentation skills (Tufte, 1983) may also be worthwhile. The aim is to make students aware of the resources that are available should they be needed, rather than to provide comprehensive coverage. Again, depending on the ideas and interests of the students, some discussions may focus on particular skills that are relevant to the researcher, such as brainstorming techniques and writing academic curriculum vitae. Finally, it is worth sketching the opposing views of science as a rational (e.g. Popperian) endeavor and as a non-rational (e.g. Kuhnian; see Kuhn, 1970) process (Newton-Smith, 1981) in order to equip students to locate their research in the context of contemporary scientific thinking.

ORGANIZATIONAL AND PERSONAL SKILLS

Students often lack the ability or aptitude to cope with having to organize themselves at university. This becomes particularly critical at the postgraduate level where they may work for days, weeks, or even months, with little external input. Simple organizational skills can make a huge difference to the amount of work that can be accomplished each day. Students also need to learn how to recognize signs of stress, and anticipate stressful situations. We find that they relate to this best if it is taught in a scientific manner; for example, a discussion of the physiological effects of stress. This is also a good opportunity for students to meet campus health-care professionals in a non-crisis situation.

SOCIAL EVENTS

Graduate programs run best if there are good personal relationships between faculty and students. A lot can be achieved by organising suitable social events. Possibilities include a lunch in the department during the first week of study, a meal at the faculty club or a local restaurant, an outing (e.g. a ski trip), a competition (e.g. an ice hockey challenge), or a party at a faculty member's house. Through a well-organised event, students can see their professors as normal people, and appreciate that they have interests outside of computer science. Also, faculty get to know a little about students' personal interests, meet their spouses or "significant others," and gain some appreciation for the pressures and demands that students face outside the university. Feedback from students indicates that social events are popular.

The two courses

The two courses were designed to meet a perceived need to accelerate students' progress at the beginning of their research program. Both draw on the material identified above, although neither covers it all. The credit course has been given for three years to incoming graduate students (primarily MSc students) in the Computer Science Department, University of Calgary, Canada; enrollment ranged from 8 to 18. The short course has been given to about 20 incoming graduate students and fourth-year Honors undergraduate students in the Computer Science Department, University of Canterbury, New Zealand. It concentrates on some of the more practical skills identified above.

THE SHORT COURSE

The short course is held during the first week of term, a time which is not always very productive anyway. As well as focusing students' minds on research from the beginning of their program, it provides an opportunity for them to learn how the department operates, meet faculty, and prepare themselves for the year ahead. It also serves a social function and ends with a party so that new students get to know each other, and other members of the department, in an informal atmosphere.

The course lasts a week. We strive to make it informative, enjoyable, and informal; and take care not to make it too demanding. We have heard of other departments offering intense short courses which have had the opposite effect to that desired: the sheer pace

has discouraged students from continuing! The lecture load leaves plenty of time for students to discuss research topics with staff, and set up their office and computing environments.

Most of the sessions for the short course are run as informal seminars, convened by a member of faculty. Table 1 shows the topics covered. Questions and discussion are encouraged. Outside help is enlisted for some of the seminars; for example, the literature search session is lead by librarians, in the library, so that students became familiar with staff and facilities there. The time and stress management seminars are given by experts from outside the department.

Because the course is so brief there is little time to practise the skills discussed. It is up to students' supervisors to encourage them to adopt the good habits that have been introduced.

THE CREDIT COURSE

For the credit course the class meets once a week for two semesters, which is half the intensity but twice the normal duration of a half-credit course; this seems appropriate for the type of material covered. The first hour of each 90-minute class is generally a full-length seminar (which all graduate students and faculty are invited to attend); this is followed by a briefer presentation or discussion period. Towards the beginning of the course the full-length seminars are generally presented by faculty; around the middle, by senior graduate students; and in the later part, by students registered in the course. However, this structure is quite flexible, particularly in accommodating the needs and desires of senior graduate students to present and obtain feedback on their work. The short presentations are generally given by the instructor or students in the course.

The organization of the course depends on the number of students enrolled. Once this exceeds 16 or 17, changes are required to accommodate all student presentations, whereas smaller numbers give more opportunity to include classes on diverse tools and techniques. Table 2 shows a typical schedule for an enrollment of eleven students. Even in this extended version of the course there is not a great deal of time to cover the topics reviewed above—only fourteen sessions, most of which are short ones. It is much better to pack information into fast-paced classes than to run the risk of having to spread

material too thinly. Much of what students learn is through practicing relevant skills rather than listening to lectures on research methods.

For example, students are expected to give two seminars to the class. One is brief, limited to 15 minutes and followed by a 5-minute question period—a common format for academic conferences. It introduces and critically assesses a published research paper as though it were being presented at a conference. The other lasts 50 minutes followed by a question period—a common format for major presentations and seminars. It focuses on a particular, narrowly-defined, topic, perhaps presenting and contrasting the approach taken in a small group of related papers. The class discusses the presentations; this is also extended to critical discussions of regular departmental seminars, in which the students are expected to participate fully.

Students gain experience with reviewing by writing two reviews in the style of *Computing Reviews*, after having studied good and bad examples. Students themselves, in conjunction with their research supervisor, select papers to review. Appropriate choices can give a student a head start on thesis work. Quite apart from the goal of honing critical evaluation and review-writing skills, students who choose papers intelligently will, by the end of the course, have read and thoroughly analyzed several key papers in their area of interest.

Students experience what it is like to examine a graduate thesis through a "mock" oral in which a senior student defends on his or her thesis to the class. This gives valuable insight into the problems that an examiner must face. The preparation of a standard examiner's report on the thesis is a useful exercise that allows students to practice their skills of critical analysis and compels them to reflect on what makes a good thesis.

To provide a practical context for teaching about how to do research we focus on research proposals as a vehicle for communicating information about good research skills (Witten, 1990). A thesis proposal is a required part of the credit course—this also helps students begin their thesis research. The proposal is judged on the idea itself, on how it is turned into a well-circumscribed research question, on the relevant background knowledge demonstrated, and on the clarity, conciseness and eloquence of the writing. It is limited to 10 single-spaced typeset pages, and clarity and presentation is paramount. Useful sources include Langley (1990), Knuth *et al.* (1988), and Sides (1984).

Assessment for the credit course is based on a number of small components, shown in Table 3. The assignments are spread fairly evenly over the course, beginning with the brief presentation and reviews and ending with the mock oral and thesis proposal.

Conclusions

Student evaluation of both courses has been generally extremely positive—apparently the existence of this course has elicited envy from graduate students in other departments! Students comment that:

- they learn less from their peers' presentations, more from faculty and senior students;
- faculty and senior students must be sensitive to the objectives of the course when designing their presentations;
- uniformity of grading is a potential problem because of bias due to the instructor's own research interests;
- organization becomes unwieldy with more than about seven to ten students;
- the course is hard work!

Also, students point out that even the credit course makes no real attempt to *teach* skills of presentation, writing, and critical analysis; it just gives the opportunity to practise. Students in the short course commented that seminars by senior graduate students about their research would be valuable—this is already a feature of the credit course.

Both courses have been successful, although it is hard to quantify the degree to which the aims have been met. They have been interesting to develop and fun to teach. The material has slowly evolved as more useful information was discovered on skills for the professional researcher. We believe that this kind of course provides a valuable means of strengthening a graduate program in computer science and enabling students to work more efficiently and productively.

References

Anonymous. The publication game: beyond quality in the search for a lengthy vitae. *J Social Behavior and Personality* 2, 1 (1987), 3–12; and comments on pp.13–22.

Brooks, F.P. *The mythical man-month.* Addison-Wesley, Reading, MA, 1978.

Bundy, A., du Boulay, B., Howe, J. and Plotkin, G. The researchers' bible. Teaching Paper No. 4, Department of Artificial Intelligence, University of Edinburgh, 1986.

Campbell, W.G. & Ballou, S.V. *Form and style: theses, reports, term papers.* Houghton Mifflin, Boston, MA, 1974.

Cleese, J. *Please can we have that the right way round?* Produced by Video Arts (20 min 16 mm film); distributed by International Telefilm Enterprises, 1976.

Crawford, S. & Stucki, L. Peer review and the changing research record. *J American Society for Information Science 41*, 3 (1990), 223–228.

Crutcher, K.A. How to succeed in science. *Perspectives in Biology and Medicine 23*, 2 (Winter 1991), 213–218.

Habeshaw, S. & Steeds, D. *53 interesting communication exercises for science students.* Technical and Educational Services Ltd, 1987.

Knuth, D.E., Larrabee, T. and Roberts, P.M. Mathematical writing. Report STAN-CS-88-1193, Department of Computer Science, Stanford University, 1988.

Kuhn, T.S. *The structure of scientific revolutions.* University of Chicago Press, 1970.

Langley, P. Advice to *Machine Learning* authors. *Machine Learning 5* (1990), 233–237.

Mental, X. Perry How to make a scientific lecture unbearable. *J Irreproducible Results 18*, 3 (March 1971).

Michaelson, H.B. *How to write and publish engineering papers and reports.* ISI Press, Philadelphia, 1982.

Newton-Smith, W.H. *The rationality of science.* Routledge/Kegan Paul, Boston, 1981.

Parberry, I. How to present a paper in theoretical computer science: a speaker's guide for students. *Sigact News 19*, 2 (Summer 1988), 42–47.

Parberry, I. A guide for new referees in theoretical computer science. *Sigact News 20*, 4 (Fall 1989), 92–109.

Sides, C.H. *How to write papers and reports about computer technology*. ISI Press, Philadelphia, 1984.

Smith, A.J. The task of the referee. *IEEE Computer 23*, 4 (April 1990), 65–71.

Strunk, W. & White, E.B. (1979) *The elements of style*. MacMillan, New York.

Tufte, E.R. *The visual display of quantitative information*. Graphics Press, CT, 1983.

Tukey, J.W. *Exploratory data analysis*. Addison-Wesley, Reading, MA, 1977.

University of Chicago Press *The Chicago manual of style (13th edition)*. Chicago, 1982.

van Leunen, M-C. *A handbook for scholars*. Alfred A. Knopf, New York, 1986.

Williams, R. *The Mac is not a typewriter*. Peachpit Press, Berkeley, CA, 1990.

Witten, I.H. How to get a research grant. *Canadian Artificial Intelligence 24* (July 1990), 9–13.

topic	hours
Choosing research projects	1
Time management	2
Stress management	1
LaTeX	2
Scientific writing	1.5
Typography	1
The literature search	2
Giving a seminar	1
The publication business	1
Socializing	party

Table 1 Topics in the short course

Brief presentation	10%
Full-length presentation (including handout)	20%
Reviews (2)	25%
Thesis proposal synopsis	5%
Thesis proposal	25%
Examiners' report and mock oral	5%
Seminar digests and participation	10%

Table 3 Assessment for the credit course

Week	Responsibility	Activity	*Assignment due*
1	Instructor	Introduction to the course	
	Head of Department	Departmental expectations of graduate students	
2	Faculty member	Research seminar	
	Instructor	Discussion of Bundy *et al.* (1986)	
3	Technical staff	Departmental computing environment	
	Instructor	Document preparation tools	
4	Faculty member	Research seminar	
	Instructor	Computer Science literature, and the library	
5	Faculty member	Research seminar	
	Movie	Giving presentations	
6	Faculty member	Research seminar	
	Instructor	Discussion of seminar presentation styles	
7	Faculty member	Research seminar	
	Instructor	Writing critical reviews	
8	Faculty member	Research seminar	
	Student	Brief presentation 1	
9	Senior grad student	Research seminar	
	Student	Brief presentation 2	
10	Students	Brief presentations 3, 4, 5	
11	Students	Brief presentations 6, 7, 8	*Review 1*
12	Students	Brief presentations 9, 10, 11	*Seminar digests*
13	Faculty member	Research seminar	
	Invited guest	Brainstorming techniques	
14	Instructor	Preparing research proposals	*Review 2*
	Invited guest	Visual display of quantitative information	
15	Students	Full-length presentations 1, 2	*Thesis proposal synopsis*
16	Students	Full-length presentations 3, 4	
17	Senior grad student	Research seminar	
	Instructor	Writing academic resumes	
18	Instructor	Discussion of students' ideas for thesis proposals. Refereeing	
19	Students	Full-length presentations 5, 6	
20	Student	Full-length presentation 7	
	Invited guest	Symbolic algebra package	
21	Students	Full-length presentations 8, 9	
22	Students	Full-length presentations 10, 11	*Thesis proposal*
23	Instructor	Thesis orals. The publication business	
24	Instructor	Mock thesis oral	*Examiner's report*
25	Instructor	Science and ethics. General discussion of course	

Table 2 Typical schedule for the credit course (11 students)