Department of Computer Science

James Cook University

Guide to Research

Edited by Xindong Wu and John Farrell

August 1994

(Last updated in February 2004)

Introduction

This **Guide to Research** was assembled from a series of presentations organized by Dr Xindong Wu and held within the Department of Computer Science at James Cook University earlier this year. These presentations were given by members of academic staff and attended by them, postgraduate and honours students. The idea behind the Guide and the presentations was to address general issues of research and research supervision from the perspective of computer science. Although there are themes common to most areas or research and supervision, there are sufficiently many issues peculiar to computer science to warrant a Guide devoted exclusively to it.

In many ways the Guide was prompted by two events external to the department. One was a workshop which I attended on postgraduate research supervision held by the Tertiary Education Institute at University of Queensland in September 1993. The other is a review of the structure of Ph.D. programs at James Cook.

I would like to remark briefly on each of the chapters making up the Guide.

Chapter 1. The Head of Department, Prof. Gopal Gupta makes several points about the overall character of research within the computing discipline. This includes mission statements at the national level and some observations about funding.

Chapter 2. Dr. Geoff Sutcliffe discusses strategies for establishing research ties with other institutions. This is especially important for James Cook because of its relative isolation, both outside and inside Australia.

Chapter 3. Dr. Xindong Wu talks about the significance of funding to research. This is a crucial component in the support of postgraduate research in computer science. Criteria used by funding bodies are listed.

Chapter 4. Dr. Tony Sloane discusses computer science writing. Particularly important are a listing of the kinds of papers and reports play a role in dissemination of results in computer science and essential components of a good conference or journal paper.

Chapter 5. Dr. John Farrell outlines major steps in designing and giving a computer science research talk. The complexity and frequent mathematical nature of much computer science research distinguishes its seminars from those of many other disciplines in important ways. This chapter is especially important because of the unique role played by conferences in computer science.

Chapter 6. Dr. Bruce Litow considers several points relating to the mutual responsibilities and expectations of supervisor and postgraduate student. Again the dynamic and complex arena of computer science offers real challenges to this already sensitive issue. Emphasis is placed on how a student arrives at a topic that can be expected to be pursued realistically to a satisfactory conclusion.

Chapter 7. Dr. John Farrell presents several interesting tips on how to use WWW (world wide web) and other information resource retrieval systems available via the InterNet. With the rise of technical report libraries and distribution of results via the Net, this topic is central to enabling researchers to discover and communicate results rapidly. It also promotes a new kind of 'prereview' process for research.

Chapter 8. Dr. Bruce Litow touches on some of the opportunities and problems connected with working outside of academia.

Chapter 9. Dr. Xindong Wu talks about the differences in acceptance criteria and the relative merits of submitting work to journals and various types of conferences.

Chapter 10. Dr. Shyam Kapur discusses the responsibilities of acting as a reviewer for journals and conferences. This is the 'unsung' but critical component of research activity.

Bruce Litow

Acting Head of Department

Table of Contents

Chapter 1

Research in General

Gopal Gupta

1.1 Background

Research recognition is primarily based on research publications in international journals and major international conferences, and research funding from external bodies. To a lesser extent, research recognition is based on consulting, publications in local and regional conferences, organisations of courses and conferences, etc.

Research performance is important since it has a direct impact on departmental funding and departmental reputation (which, for example, affects recruiting). Research performance is also important in that it has indirect impact on the Department's well being. It afffects how the University (that is, the VC, PVC) view the Department. This in turn determines the influence that the Department has in the University's decision making processes.

In discussing a research agenda for the Department we should take note of recent publications and studies like the Discipline Review of Computing Studies, Computing the Future, etc. We should be aware that quite a lot of research funding that governments hand out is not for basic research but for research that can demonstrate positive benefit for the nation. This does not necessarily mean that we should do no basic research but any basic research that we do do ought to be relevant.

In selecting research topics, we should not ignore industry trends, for example the strong trend to smaller and portable systems that are connected by networks. In software, there is likely to be an increasing trend to much more friendly user interfaces that will use multimedia, graphics and visualization. We must take these things into account in our research if we are not to become irrelevant.

1.2 Computing the Future

Some conclusions of this study:

1. CS has established an unique paradigm of inquiry that is applicable to wide variety of problems.

2. Distinctions between basic and artificial research in CS are artificial.

3. Growing ubiquity of computing within the community places a premium on the largest diffusion of CS expertise to all endeavors of society.

Some comments in the report:

1. CS must increase its contact and intellectual interchange with other disciplines.

2. Increase the number and quality of applications in areas of economic, commercial and social significance.

3. Increase traffic in CS-related knowledge and problems among academia, industry and society at large.

4. Increase the cross-fertilization of ideas between theoretical and experimental work.

Recommendation 1: The HPCC program should be fully supported.

Recommendation 2: The Federal Government should initiate an effort to support interdisciplinary and applications oriented CS&E research in academia.

Recommendation 3: Academic CS&E should broaden its research horizons, embracing as legitimate and cogent not just research in core areas but also research in problem domains that derive from nonroutine computer applications involving substantive and intellectually challenging problems.

Recommendation 4: Universities should support CS as a laboratory discipline.

Recommendation 5: HPCC program should be expanded to address educational needs.

Recommendation 6: CS&E departments should require PhD students to take a graduate minor in a non-CS field.

Recommendation 7: The academic CS&E community must reach out to women and minorities.

1.3 Discipline Review

Federal government funding is the largest source of funding for IT research in the higher education sector. Support for major centres of research is provided through the Special Research Centres, Key Centres for Teaching and Research, and Cooperative Research Centres programs. Currently there are five Key Centres, one Special Research Centre and one Cooperative Research Centre in IT (??). Funding for industry related R&D in IT is provided through the Grants for Industrial Research and Development (GIRD) Generic Technology Grants Scheme. Large grants from the Australian Research Council (ARC) are the main source of direct funding for curiosity based research in IT in higher education institutions. Funding by industry of IT research in institutions is small, with only six departments receiving more than $500,000 industry funding per annum, and only 13 more than $100,000.

1.3.1 ARC Funding of IT

The ARC does not fund IT as a distinct field of study. There is no specific funding allocation and no specialist panel to adjudicate IT applications. IT is a relatively new discipline with relatively small volume of articles published in journals compared with mathematics. The Committee notes that ARC funding for IT has risen in recent years. In rolling grants process, adjustments in the relative funding of different research areas are necessarily gradual.

1.3.2 Mission Oriented Research

In Australia, in general, the research has attracted more funding and more research students when it has been mission oriented (whether strategic or applied) - directed toward solving problems with foreseeable, perhaps long term, applications - rather than curiosity based. Mission oriented research is not necessarily confined to IT engineering: it often generates problems which can best be solved by abstraction into the science base. The attitudes of both government and industry strongly suggest that mission oriented research will continue to attract the overwhelming majority of funding.

Even the ARC, the sole potential source of support for curiosity based research, awards many of its grants to mission oriented projects, perhaps because the better applications have that bias. In any event, the apparent better funding of mission oriented projects needs to be recognised by academic departments seeking to establish or expand their research activities. The existence of external funding for research can assist in building Honours and postgraduate numbers and significant postgraduate numbers can also provide indirect assistance for undergraduate teaching.

Institutions wishing to develop or expand a research base should focus initially on mission oriented research (strategic or applied) and seek industry involvement.

1.3.3 Industry Involvement with Research

In Australia there is little interaction between industry and government funded research activity outside the agricultural and mining sectors. As a consequence there is little private sector funding for mission oriented research in IT (strategic or applied) and almost none for curiosity based research.

Australia's IT producing sector is small and is dominated by multinational corporations with little incentive to undertake R&D far from their centres of management. The expenditure that does occur tends to be confined to the final stages of product development. There is little interest in, or understanding of, the innovative role that can be played by academics. On the other hand, there is little understanding within institutions of the processes by which an innovation becomes a commercial reality, or the amount of expenditure required in developing a new product.

A few quality IT research programs have been established in Australian universities in recent years. However, very little of this research has contributed to the development of computer systems and products. There is considerable potential for an expanded IT research program in higher education to contribute to the development of IT in Australia.

The realisation of this potential will require closer and more extensive interaction and cooperation between the universities and the IT industry, including both IT suppliers and users. Such relationships will foster joint efforts on short, medium and long term research and tasks. For example, Telecom and OTC have both forged useful relationships with a number of university research groups.

The GIRD scheme has the significant feature that it is directed towards creating more effective relationships and understanding between industry and research institutions and universities. No project can be approved without a commitment of a portion of the necessary funds by the sponsoring industrial organisation. Progress has been steady, but slow, but the scheme has been successful in building positive attitudes and increased IT R&D activity in Australia.

The Partnerships for Development Program has been successful in encouraging multinational companies to undertake R&D and export oriented activities in Australia, both through their subsidiaries and through collaboration with Australian companies and research establishments, including the higher education institutions.

The Partnerships for Development Program is a key element of a national strategy initiated in September 1987 for the development of the Australian Information Industries. Partnership agreements provide an alternative longer term arrangement for multinational companies that do not wish to participate in the Australian Civil Offsets Program.

A multinational company signing the standard partnerships agreement with the government commits to achieving by the seventh year:

1. expenditure on R&D in Australia equivalent to five per cent of its annual local turnover;

2. exports of goods or services equivalent to 50 per cent of its annual imports into Australia; and

3. an average of 70 per cent local value added across all exports.

Under some of the early agreements, multinationals agreed to reach their targets in less than seven years. The government undertakes, among other things initiated as part of the Information Industries Strategy, to take steps to ensure there is an adequate supply of skilled labour and to establish a Vendor Qualification Program. While the achievement so far has been significant, there is considerable potential to achieve much more of national economic, educational and social benefit by continuing expansion of the program.

1.4 Aims of a Research Agenda

- All researchers benefit from it in terms of their professional development and marketability of their skills.
- The research conducted in the Department is relevant to the needs of the country, region and the University.
- No researcher is involved in more than two active research areas.
- Each researcher is involved in some mission-oriented research.
- Each researcher is working in at least one group of more than two researchers.
- Research activity - there should be no research activity in the Department that is not known to the HOD. This includes all activities other than that involved in normal teaching, including writing of survey papers, books, consulting, etc.
- Research publications - there should be no paper submitted to any journal or conference or workshop without prior permission of the HOD.
- Research travel funding

Chapter 2

Interacting with Researchers from Other Institutions

Geoff Sutcliffe

2.1 Why do it?

There is not enough interaction here

There is not enough expertise here

2.2 How to establish contacts

Go to conferences

Respond to articles on the news

Write personal letters

2.3 Who will respond?

People with time to help and deep interest

People who can gain from helping you

2.4 The importance of the Internet

Must have good Internet access

Must use the Internet continually

2.5 Co-operating

Reply to email fast (modulo time zones)

Keep old email intelligently

Include email in replies

Establish protocols for joint authoring

2.6 Interaction

Offer something

Reply to email fast, and always keep the door open

Visit and give seminars

Write joint papers

2.7 Acknowledgements

Do it in anything you produce

Send copies of your work

Chapter 3

Funding Applications

Xindong Wu

This document relates to the following three questions:

1. Funding: How important is it in research?

2. Funding channels

3. Writing up a good application

3.1 Funding: How important is it in research

Research recognition:

- Refereed publications in journals and conferences

- Research funding from external bodies

- Involvement in international academic activities - your standing in your research areas, not just service

- Awards - society awards and best paper awards

- Invited/keynote conference talks

- Citation of your research by others
- Use of your software products (not just for money)

Research funding can help both your publications and your involvement in international academic activities in a number of ways.

With sufficient research funding, you can employ research assistants (including research students), buy more advanced facilities, and improve the research profile of both yourself and your institution.

3.2 Funding channels

- University scholarships and research fellowships (with ARC, for example)
- Basic/curiosity based research (ARC Large and Collaborative grants) – Potential application of research results is always worth your while to mention.
- Mission oriented (strategic or applied) grants (DEET Evaluation and Investigation Program, Grants for Industrial Research and Development (GIRD) Generic Technology Grants Scheme)
- National and international collaboration

3.3 Writing up a good application

Criteria for application assessment:

- Is the proposed project interesting/promising?
- Does the chief investigator (CI) have a good background?
- Is the CI suitable to carry out the proposed project?

The applicant's research background is extremely important. In theory, the funding panel cannot reject an excellent proposal, but in reality, your application still cannot be granted without the support of good track records.

Good points:

- Read the application forms and make sure that what you are proposing is interesting to the finding body. Otherwise it would be a waste of your time to write up an application.
- Make your aims and significance very clear.
- Say what you have done as well as what you are going to do.
- Cite good references to show that you are knowledgeable in your field.
- Be reasonable with your budget.
- Be careful in choosing collaborators.

Chapter 4

Computer Science Writing

Tony Sloane

4.1 Computer Science Writing

- Write to explain and explore your ideas and make them accessible to others.
- Types of writing
  - technical reports
  - workshop or conference papers
  - journal articles
  - thesis proposals, theses
  - grant proposals, progress and completion reports
  - manuals, users guides

4.2 Resources

- General writing:
  - Strunk and White: *The Elements of Style*
  - Shertzer: *The Elements of Grammar*
  - Gowers: *The Complete Plain Words*
- Scientific writing:
  - Markel: *Technical Writing: Situations and Strategies*
  - Davis and Parker: *Writing the Doctoral Dissertation*
  - Lauer: *On Ph.D. Thesis Proposals in Computing Science*

4.3 The Writing Process

1. *Assess the audience:* To whom are you writing? Why will they be reading your writing?
2. *Assess the purpose:* Why is it being written? What should readers take away?
3. *Determine an overall strategy:* Scope, structure, organisation, tone and vocabulary.
4. *Brainstorming:* 15-20 minutes of writing down as many ideas about your subject as you can (in no particular order).
5. *Outlining:* Put similar ideas into groups, sequencing within groups and sequence groups. Fill gaps.

6. *Drafting:* Turn your outline into paragraphs.

7. *Revising:* Read and reread your draft with attention to content and organisation.

8. *Testing:* Give it to someone who is as close as possible a “typical” audience member. Address their comments.

4.4 The Importance of Reading

- Reading other people’s writing provides a good source of examples of how to write (and how not to write).

- Ask others for pointers to papers that they think are well-written.

- Can often provide useful guidance on the best way to organise and present information.

4.5 Know Your Enemy

- For conferences and journals, read published papers in those forums. Find out who’s on the program committee or editor board.

- For theses, read previous theses from the same institution. Learn what your advisor likes.

- Make sure you know the publication guidelines (length limits, formatting standards etc.)

- Read as many “How to” papers as you can, particularly ones directly applicable to your target.

4.6 Originality

- Are the ideas in the paper new?
- How do you know?
- Can you state the new idea concisely?
- What exactly is the problem being solved?
- Are the ideas significant enough to justify a paper?
- Is all related work referenced, and have you actually read the cited material?
- Are comparisons with previous work clear and explicit?

4.7 Reality

- Does the paper describe something that has actually been implemented?
- If the system has been implemented, how has it been used? What did this show?
- If the system hasn't been implemented, do the ideas justify publication now?

4.8 Lessons and Choices

- What have you learned from the work?
- What should the reader learn from the paper?
- How generally applicable are these lessons?
- What were the alternatives considered at various points, and why were the choices made the way they were?
- Did the choices turn out to be right, and if so, was it for the reasons that motivated them in the first place? If not, what lessons have you learned from the experience?

4.9 Context and Focus

- What are the assumptions on which the work is based?
- Are they realistic?
- How sensitive are the results to perturbations of the assumptions?
- Does the introductory material contain excess baggage not needed for your main development?
- Do you include just enough material from previously published works to enable your reader to follow your arguments?

4.10 Presentation and Writing Style

- Are the ideas organised and presented in a clear and logical way? Is the writing style clear and concise?
- Are terms defined before they are used?
- Are forward references kept to a minimum?
- Does the abstract communicate the important ideas in the paper?
- Are words spelled and used correctly? Are the sentences grammatically correct?
- Are ambiguity, slang and cuteness avoided?

Chapter 5

What to Say in a Good Research Talk

John Farrell

Giving a good research talk is not easy. Hints on mechanical aspects of the speaking process, such as format of accompanying overheads, behaviour during the talk, and so on, abound, but few authors focus on the content of the talk; i.e. on the process of designing a good talk in the first place. I make some suggestions to assist novice and experienced researchers alike to construct a good research talk. The structure of this document (though not the content) is heavily based on the paper by Simon Peyton Jones, John Hughes and John Launchbury. That is a useful paper, and I recommend that it be read as a companion to this one.

5.1 What This Document is About

By a research talk, I mean a presentation of some research work you have done. This work is original to you, and hopefully original to your audience as well. Thus my basic premise is that you know something that the audience doesn't, and the purpose of the talk is to tell them about it. My hypothesis is that if you start by saying the right thing, finish by saying the right thing, and say all the right things in the middle, you have given a good talk. Consequently, my guidance is based on choosing how to start, how to finish, and what to do in the middle.

5.2 Where to Start

You must start your talk with something the audience can understand. In a research talk, this is usually a statement of the problem. Exactly what you choose as your problem statement depends on your audience.

Generally speaking, your audience will consist of the following types of people:

1. Graduate Students

2. Scientists who do not share your research interests

3. Scientists who do share your research interests

Graduate students, depending on the advancement of their studies, probably don't know much about the forefront of research in your chosen area, nor even a lot of the background, so to speak to them it would be necessary to identify a well-known problem (we need to cure cancer, we need to write software that works). Scientists who do not share your research interests are likely to be familiar with the important problems in the area, so you are able to choose a less general problem to start with (we need to discover the structure of a virus, we need to make programs more readable). Finally, scientists who do share your research interests not only understand the problem you are trying to solve, but will be familiar with previous attempts, will have ideas of their own, and may be conducting similar research. In this case, you may choose a very specialised statement of your problem (we need to determine when eager evaluation may be substituted for lazy evaluation without introducing non-termination).

The composition of your audience will depend on the forum - a departmental seminar is likely to consist mostly of types (1) and (2), an international conference will be mostly types (2) and (3), and a specialised workshop will be mostly type (3). In each case, you should choose a different problem statement, but in all cases it should be a problem that the audience already appreciates to be a problem. If they agree with your choice of problem, they will be more willing to listen to your method of solution. If you start by saying something the audience doesn't believe, they won't believe the rest of what you say either.

Holtzman [1970] calls the starting point on which everyone agrees the *Common Premise*. The more specific your Common Premise is, the less ground you have to cover to your solution, and the more time you will have to cover your work in between. Your Common Premise should understood by the majority of the audience, and believable with a quick explanation to the others. This gives us the following rule to choose where to start the talk:

Rule 1: Choose the most specific Common Premise suitable to the audience.

5.3 Where to Finish

The preponderant tension in construction of a seminar is between the small amount of time available and the large amount that you could say about your topic. No matter how fast you speak or how many overheads you prepare, you will only be able to communicate a limited number of ideas in the time period, so you should aim to present only that many ideas. If you have written a paper to go with the talk, put everything you want to say in the paper; if your talk is successful people will read it anyway. So, as you can't say everything you want to, you must carefully choose the things you do say.

My technique to achieve this is to choose the single most important thing I want the audience to learn, and construct the rest of the talk so as to achieve this single goal. I shall call this goal the Ultimate Idea. The Ultimate Idea should be characteristic of your work, with the result that comprehension of the Ultimate Idea gives a listener an understanding of your research.

The Ultimate Idea is not necessarily, and in most cases hopefully not, a summary of the entire field. Any field which deserves devoted research is too complex to be summarised in a short talk, and you should not aim to do so. However, to keep the attention of the audience, the Ultimate Idea should be novel, unusual or unexpected, and it should be a solution to, or related to a solution to the Common Premise. It should be an idea which demonstrates the benefit of research in your field, and your aptitude at carrying it out. We have the following rule to choose where to finish the talk:

Rule 2: Choose an Ultimate Idea which characterises your work, and demonstrates progress towards the solution of the Common Premise.

5.4 Where to Go in Between

Having chosen how to start the talk, and how to finish it, the structure of the body of the talk becomes obvious - you present ideas in a sequence which leads the audience naturally and comprehensibly from the Common Premise to the Ultimate Idea. The most important restriction on this progress is:

Rule 3: Proceed directly from the Common Premise to the Ultimate Idea. Do not get sidetracked.

Irrelevant sidetracks will confuse your audience and lose their attention. If you can't bear to leave a sidetrack out, it is worth presenting in a separate seminar, but don't clutter up this one!

The key to presenting successful research talks is to keep it simple, even painfully simple. The people you are speaking to do not have experience with your ideas and notation, and will not comprehend your ideas instantly. However if you can present a series of clear steps which lead from a problem to a solution, your audience will understand and you will have achieved your goal in presenting the seminar.

We now present a series of ideas which help you achieve comprehension and credibility.

5.4.1 Assert Your Authority

It is important to assert your authority to talk on the topic, as listeners not familiar with your work will be wondering whether you know anything or not. One way to demonstrate your qualifications is to give a concise and accurate summary of work done on this particular problem. This may be combined with the Common Premise, by giving a history of work leading from the general problem to the specific problem that you are addressing. You may demonstrate that you are familiar with the field by outlining major developments, and accredit them to known authors in the field. This "namedropping" insinuates that you are so familiar with the field that you almost personally know other researchers.

Of course, if you are a well-known researcher in the field, either your audience will know you already or you can start your talk with a history of your own work in the area ("since I proved the Continuum Hypothesis, I have been investigating...").

5.4.2 Tell the Truth

In a court of law, witnesses are required to tell "the truth, the whole truth, and nothing but the truth". This is not the case in a technical seminar, as your listeners are not being paid to hear you speak. The whole truth is not necessary.

However it is often the case when presenting technical research that the reasons behind assumptions, proofs, etc, are very complex and require a great deal of explanation. This explanation should go into an accompanying paper, not into your talk. If you have successfully established your credibility with the audience, you can use up some of it by asking them to believe you. However this is bad practice and should be avoided if at all practical.

It can be sufficient simply to give an intuitive grasp of the reasons behind your dubious statement (this is called "hand-waving"). This can be done by mentioning a related proof which is well-known, or by demonstrating a crucial part of the proof.

If you can't convincingly avoid this proof, it may be necessary to reconsider the design of your talk. If the proof is so important, it should probably have a talk of its own.

Returning to the court room, the requirements that you tell the truth and nothing but the truth remain. You are giving evidence to an audience which probably contains experts, and they will be less interested if you do not tell the truth. Furthermore, they could expose you as a charlatan during question time.

Rule 4: Tell the truth, but not the whole truth.

5.4.3 A Picture's Worth a Thousand Words

A more fancy and effective method of hand-waving is by drawing a picture. If an audience can see a graphical demonstration of why some statement is true, they are likely to believe you quickly. Furthermore, they are likely to understand the picture far more easily than they are a proof.

Rule 5: A picture's worth a thousand words, and billions and billions of Greek letters.

5.5 After the Talk

Any good talk deserves questions, and the lack of them is a sure sign of failure. If you do know what you are talking about, questions are to be welcomed rather than feared. As Lethbridge [1991] points out, most questions will be of a curious and sympathetic nature, and demonstrate interest in your talk. You should answer them with enthusiasm; the questioner may turn out to be a valuable colleague.

We now examine some common types of questions, and techniques for responding to them.

5.5.1 Questions about Material You Covered

Some questions will be about material you did cover in the seminar - the questioner may have misunderstood or misinterpreted what you were saying. Maybe you covered it too fast, maybe the questioner has it confused with some similar work.

In any case, this is the sort of question you should be able to answer - you chose to put the material in your talk, you should be able to explain it. If only one listener has such a problem, you shouldn't worry. If the whole audience needs clarification, you should rethink that part of the talk before you give it again.

5.5.2 Questions about Material You Didn't Cover

Many questions, and in some ways the most flattering, will be requests for additional detail on areas you didn't cover due to time constraints. However do not take this sort of curious question as an invitation to present another five slides!

If the question probes an area of which you do have a wide knowledge, say enough to satisfy the questioner but not enough to bore the audience as a whole. You might like to talk to the questioner later.

If the question probes an area which you are unfamiliar with, it is your job as a reputable researcher to have a reason for not knowing. A standard line is "I haven't got around to looking at that yet."

5.5.3 Doubts about Your Common Premise

Some members of the audience may have doubts about or serious disagreement with your common premise. However relevant their doubts are, your seminar is not the place to start a debate on the matter - in your seminar, the Common Premise is taken to be true! With experience in the field, you will learn the ways in which your Common Premises are attacked, and you will have to develop ways to answer these questions. While such questions are not welcome in your seminar, it is your duty as a researcher to consider their merit, so that you remain convinced that your research is worthwhile.

5.5.4 Irrelevant Rambling

Occasionally a questioner will take the opportunity of your question time to sermonise on a marginally relevant topic. In this case, it is often difficult to discern the question, let alone answer it. Such a speaker must be diplomatically shut up - the audience is here to listen to you, not them! Depending on your composure when faced with invasion of your fifteen minutes of fame, you may promise to "discuss it later", or give a short sermon of your own to assert your occupation of the pulpit.

5.6 Conclusion

In summary, you must design the talk for your audience. Start the talk with something they know, finish it with something they didn't know. In the middle, say things they understand. At all times, remember that you are, and act as if you are, an expert in your field.

5.7 References

[**Holtzman 70**] Paul D. Holtzman, *"The Psychology of Speakers' Audiences"*, Scott, Foresman's College Speech Series, Scott, Foresman and Company, 1970.

[**Lethbridge 91**] Roger Lethbridge, *"Techniques for Successful Seminars and Poster Presentations"*, Longman Cheshire 1991, ISBN 0582869838.

[**Peyton Jones, Hughes & Launchbury 93**] Simon L. Peyton Jones, John Hughes & John Launchbury, *"How to Give a Good Research Talk"*, ACM SIGPLAN Notices, Volume 28, No. 11, November 1993.

Chapter 6

Postgraduate Research Supervision

Bruce Litow

This workshop will focus on just one issue:

CHOICE OF PROBLEM

Why is this so important?

6.1 Checklist 1: Problem Choice

1. Who has priority in the choice, the student or the supervisor?

2. To what extent can the student modify the problem within the following time constraints?

    - Within the first six months
    - After the literature review
    - After the first year
    - Later than first year

3. Same question, but for the supervisor:

6.2 Must the Problem Be Drawn Directly from the Supervisor's Current Research Topic?

List advantages and disadvantages for both student and supervisor:

1. Problem stems directly from supervisor's research

2. Problem is in a related area, but not a current focus of the supervisor

3. Problem is not too closely related to main thrust of supervisor's research, but student has already done some work on the problem

4. Both supervisor and student are interested in the problem, but neither has investigated it

5. Can you identify another situation where supervision is possible?

6.3 How Much Literature Is Enough for a Dissertation to Get Under Way?

1. Is the problem already specified in the literature?

2. Are there partial results?

   - in the literature
   - due to supervisor
   - due to student

3. How does one decide relevancy?

4. What are the responsibilities of the supervisor in literature search?

5. Of the student?

6. For the last two points: do these responsibilities change over time?

Chapter 7

Introduction to the World Wide Web for Scientific Resource Discovery

John Farrell

When I was invited to give this talk, the suggestion was that it should be about using ftp, gopher, etc, to get scientific papers. It may seem then that I have chosen to talk about my "hobby horse", rather than the suggested topic. As you shall see, this is not the case.

7.1 What's the World Wide Web?

Those of you who are fans of cyberpunk novels will be familiar with the concept of CyberSpace - the world of information which exists parallel to our own. Ten years ago cyberspace was a fantasy. Today it exists, and it is called the World Wide Web.

To describe the web is an almost impossible task - it's too big for mere words to fit around. **Awesome** is a starting point. The web is a world-wide hypertext system integrating all public information services existing on the Internet. I can say this without fear of gross error, because anything which isn't integrated into the Web yet can be within days. But I don't know of anything significant which isn't integrated.

7.1.1 More Information

Planet Earth WWW Page List(`http://white.nosc.mil/www.html`) A list of pages describing the web.

7.2 Jargon

The web comes complete with a bewildering collection of jargon to confuse the novice. Any system worth its salt would. Here are some terms which may be confusing:

WWW, W3 : Simply means "World Wide Web". No special meanings yet.

Surfing : For some reason the term for wandering idly through cyberspace has come to be surfing. The implication is that cyberspace navigation is a task which you must concentrate on to avoid being drowned in a sea of information - there is some truth in this.

Web Walking : Web walking is automatic navigation of the net by a program. Because programs are faster and more persistent than humans, and consequently use more resources, this can sometimes be a problem. See also **World Wide Web Wanderers, Spiders and Robots** (`http://web.nexor.co.uk/mak/doc/robots/robots.html`).

URL : A URL is a **Uniform Resource Locator**, which means it is a way of describing where a document can be found, no matter where it is. These are strings taking the general form `protocol://info`, where the format of the `info` depends on the `protocol`. Protocols are described further below. The most common protocols are:

```
http://machine:port/data
ftp://machine/filename
file://machine/filename
```

`gopher://machine/filename`

There are examples of these below.

link : A link is a spot where a user gains access to another document through the hypertext system. On the Web, links may be connected to text, on images, on parts of images.

HTML : HTML (HyperText Markup Language) is a language based on SGML which is used to write hypertext. Its most important feature is the ability to include URLs as links to other documents. A new version of HTML called HTML+ is being designed.

CERN : A nuclear physics research centre in Switzerland where the Web was invented.

NCSA : The National Center for Supercomputing Applications, an instition in the USA which produces popular Web software.

7.3 Protocols, Servers and Browsers

It is necessary to make clear the distinction between protocols, servers, and browsers, as casual users of the Internet may not be entirely familiar with the distinction.

- A *protocol* is a data format by which information is transferred.
- A *browser* is a program with which users can cause data to be transferred.
- A *server* is a program (often a daemon) which is used to provide information in response to requests from a browser.

The concept of `ftp` unifies all these into one confusing mass - a user starts the ftp browser, which connects to an ftp daemon on another machine, to arrange a file transfer using the ftp protocol. Some common protocols are:

FTP : A TCP/IP application level protocol for file transfer.

TELNET : A TCP/IP application level protocol for interactive sessions. Gopher]: A protocol for text-based directory structured information browsing and retrieval.

Z39.50 : A protocol for document retrieval by keyword. Implemented by WAIS (Wide Area Information Server).

HTTP : A very general protocol for retrieval of arbitrary documents.

NNTP : The news protocol.

SMTP : The mail protocol.

Web browsers such as NCSA Mosaic can use *all* of these protocols except SMTP. As the Web is a very public type of system, a private system like mail is difficult to integrate. Each protocol has its own URL format.

```
file://coral.cs.jcu.edu.au/users/other/ftp/web/cs/home.html
telnet://database.carl.org/
gopher://gopher.niaid.nih.gov/77/deskref/.Dictionary/enquire
ftp://alex.sp.cs.cmu.edu/alex
http://www.sti.nasa.gov/nasa-thesaurus.html
```

7.3.1 More Information

Entering the World-Wide Web: A Guide to Cyberspace : (`http://www.hcc.hawaii.edu/guide/www.guide.html`)] A similar paper to this one, more introductory, with sections on Mosaic. Includes a glossary.

7.4 What's on the Web?

The resources on the web are those which people have so far found useful to put there. And considering it really is very useful, and it's not very difficult to add, there's an awful lot out there. The web is mind-bogglingly big already, and is growing fast. Matthew Gray has written an amusing account (`http://www.mit.edu:8001/afs/sipb/user/mkgray/ht/wow-its-big.html`) of this. Here is a list of some of the most unusual things I have found:

- Esperanto Hypercourse (`http://utis179.cs.utwente.nl:8001/esperanto/hypercourse/`)
- OFFICIAL World Cup Soccer Server (`http://sunsite.sut.ac.jp/wc94/`) - No joke, this is official. The World Cup people decided that the Internet was important enough to broadcast soccer information on.
- "The Hacker Crackdown" by Bruce Sterling (`http://www.scrg.cs.tcd.ie/scrg/u/bos/hacker/hacker.html`) - The first (but not the only) novel published on the web.
- The Exploratorium (`http://www.exploratorium.edu/`) - An interactive museum in San Francisco.

7.5 Resource Discovery

Resource discovery is a term coined by information retrieval scientists to describe the search for information on the net. It is the fundamental problem of the web today, and will remain so. It is made particularly difficult by the anarchic nature of the web - resources might be available, but they are not necessarily advertised to all potential users.

Of course, indexing is the answer, but there are several difficulties:

1. You can't index what you can't find.

2. You can't index everything, it's just too big.

The solutions to this effectively involve indexes piled on top of indexes. Some of the indexes are searchable, and some are lists of links with hints as to what might be behind them. It is impossible to know which is the best index - the web changes like a living creature and pointers into it are always out of date, and never comprehensive.

The art then of efficient resource discovery involves knowing which indexes to look at: which to search and which to investigate for links. These skills are effectively acquired by experience, rather like searching for information in a library. The best a tutorial like this can do is give some starting points and let the user discover the rest. **Planet Earth**, mentioned above, provides links to most of the useful pages, so start there. Eventually you will compile a list of links that you find useful.

7.6 Using the World Wide Web

You should be able to run `xmosaic` on an XTerminal, or `lynx` on a text terminal. You will want to set your home page:

```
setenv WWW_HOME file://coral.cs.jcu.edu.au/users/other/ftp/web/cs/home.html
```

After that, you should be able to learn by clicking.

Chapter 8

Working outside Academia

Bruce Litow

8.1 Three Avenues

- Research Labs, e.g., ATT, IBM, Phillips
- Software
- Commercial

8.2 Software

CS graduates compete with Physics and Mathematics graduates.

Product development requires Capital. Volatility is induced by

(a) programmers

(b) market

Even size isn't a guarantee of stability

8.3 Commercial

Projects within a commercial enterprise are usually of two kinds:

(a) maintenance and tinkering

(b) installation of a package (including OS, and TP software)

It is very rare for an outfit to commit money to internally developed software.

8.4 Systems/Applications

Applications are tied to business operations. This usually means that projects are supervised from two sides:

(a) programming, typically a programmer who is now a middle manager

(b) managers, generally without DP background, or a long way from it. They are interested in business operations.

8.5 Systems

Less supervised from either applications or business operations. Liaison with OS, TP software vendors. Generally regarded with suspicion by managers and programmers alike.

Chapter 9

Submissions to Journals and Conferences

Xindong Wu

This document contains information about:

1. Why are journal articles different from conference papers?
2. In Computer Science, conference publications are as important as journal publications.
3. My personal advice on paper submissions.

The article on "Evaluating Computer Scientists and Engineers for Promotion and Tenure" (`http://www.cra.org/reports/tenure_review.html`) by David Patterson, Lawrence Snyder, and Jeffrey Ullman, is worth your while to read. This article was approved by the Computing Research Association (CRA) Board of Directors in August 1999, and can therefore be viewed as an official document.

9.1 Why journal articles differ from conference papers

- Conference papers have page restrictions - Your ideas cannot be presented in every detail.
- Most conferences don't have sufficient time to choose appropriate referees - Conference referees can be less reliable.
- The referees cannot work on conference submissions in enough detail - They are asked to read a number of papers within a limited time.
- You can still publish your idea in a journal even if it has been published in a conference.

9.2 In CS, conference publications are important

- If you work in theoretical CS, all conferences in your areas are possibly either non-refereeed at all or abstract-refereed only, and conference publications really do not mean much in this case.
- If you work in a "traditional" CS area, such as AI, databases, and software engineering, you would have experienced that getting accepted by a top conference has been *much harder* than getting accepted by a first-rate journal! The reviewers for top conferences are all carefully selected and are by and large more reliable than journal reviewers.
- If your research areas do have an applied nature, and you have never published in any top conference, people in your areas will doubt about your ability in publishing at top conferences, even if you have published tons of journal papers.

- Conference papers have page limits, but they can still be significantly longer than the "letters" published in Nature. The length of papers does not always correspond to quality, and most (if not all) first-rate journals publish short/concise papers.

- The turn-around time of conferences is significantly shorter than journals (at least on average), because conferences always have predefined deadlines.

- Some conferences like ICML have area chairs and provide conditional acceptance for minor revisions. If your paper requires a major revision, you would be rejected from a top-tier conference, and can only go to a different forum.

9.3 Advice

1. Do good work and **publish** it in time. Without good publications, you don't have much say in your field even if you hold a senior title.

2. Publish preliminary and in progress work in conferences and workshops - Don't feel ashamed if your papers turn out to be incomplete.

3. When the work matures, publish it at top-tier conferences as well as in first-rate journals.

4. The number of your publications is less important than the quality of your representative publications, but having 2-3 times the average will make you stand out.

5. Your paper can never be 100% satisfactory to anyone (including yourself!), so don't keep it until others publish the same idea first!

6. Appropriate citation of your own work normally helps.

Chapter 10

Refereeing a Paper

Shyam Kapur

10.1 Introduction

Peer review is an integral part of academic life—whether it be when one seeks employment, tenure, promotion, appointment to various positions in professional organisations etc. To alleviate the shock of some harsh realities that I believe one inevitably comes across in real life, I feel that learning from each other's shared experiences is invaluable.

By discussing refereeing, we are not only seeking to learn how to referee, but indirectly also how to write a paper for which positive referee reports are quite likely, how to choose journals/editors to send papers to, how to choose whether to send papers to conferences or journals or both etc.

10.2 Kinds of Questions Discussed

- How is refereeing different from reading a published paper?
- How much effort is reasonably expected of a referee?
- Do the referees mostly do their job well?
- How are the editors supposed to use the referee reports? How do they actually use them?
- How do conference program committees work? How much does their composition matter?

10.3 References

- The Task of the Referee—Alan Jay Smith, IEEE Spectrum, 1990.
- A Guide for New Referees—Ian Parberry, SIGACT News, 1989.

10.4 Outline of the Refereeing Process

- Editor's request
- The reply
- The reminder
- The report
- Resubmission
- Second review

10.5 Evaluating a Research Paper

- What is the purpose of the paper?
- Is the paper appropriate?
- Is the goal significant?
- Is the method of approach valid?
- Is the actual execution of the research correct?
- Are the correct conclusions drawn from the results?
- Is the presentation satisfactory?
- What did you learn?

Desirable attributes of a paper include *correctness*,*significance*, *innovation*, *interest*, *timeliness*, *succinctness*, *accessibility*, *elegance*, *readability*, *style*, and *polish*.

10.6 Referee's Report: The Cover Letter

- Your name and address
- The title and authors of the paper, and any bureaucratic code or number
- Your level of expertise in the specific subject area
- Your level of effort
- A brief summary of your recommendation and justification
- Any other confidential remarks

10.7 Referee's Report: The Formal Part

- A brief synopsis of the paper and its significance

- (Optionally) Your recommendation

- Constructive criticism of the results, proofs, and presentation

- Suggestions for possible improvement

- A table of typographical and other minor errors

10.8 Sample Conference Referee Report Form

```
Please answer the following questions by putting a number between
1 (definite no/low) and 5 (definite yes/high) in the box provided.

 - How well does the paper fit into the scope of AII'94
   as a learning theory conference?                               [ ]

 - How well does the paper fit into the scope of AII'94
   according to its special intention to be mainly focussed
   on program synthesis?                                          [ ]

Please, evaluate the paper in detail according to the following
aspects:

 - Significance:                        [ ]
 - Originality:                         [ ]
 - Validity:                            [ ]
 - Quality of Presentation:             [ ]

 - Overall Evaluation:                                            [ ]

Please, indicate your confidence
(will NOT appear in the final review)   [ ]

Recommendations to the author(s):

Justification:
(This is confidential and will NOT be mailed to the author(s))
```