

# Getting a PhD (in Computer Science) - Doing Good Research

Frans Coenen

December 2007

BCS SGAI

Cambridge

Department of Computer Science  
The University of Liverpool  
f.coenen@liverpool.ac.uk

## OVERVIEW

1. What is a PhD?
2. Deciding on a suitable topic
3. Conducting your research (identifying a methodology)
4. Monitoring progress (goals and targets)
5. Writing up
6. Are publications important?
7. Trouble Shooting
8. Summary and conclusions

## WHAT IS A PhD?

- A PhD (Doctor of Philosophy) is a piece of focussed research resulting in a thesis.
  - The result should not be a text book!
  - The result should not be a diary of events!
- 

## What do we mean by focused?

- The work is directed at a particular end goal, i.e. the establishment of a research question in a particular research area (*domain*).
  - It is not a collection of unrelated or poorly related investigations.
- 

## What is research?

- Research is “*original investigation undertaken in order to gain knowledge and understanding*”.<sup>1</sup>
- The process of research implies a systematic approach.
- It is not invention!

There is a difference!

## Systematic?

- A systematic approach implies the adoption of some methodology to arrive at the desired end goal.
- There are many treatises on research methodologies (Dawson).
- Invention, however, is founded on flashes of inspiration not rooted in a systematic research methodology.
- The danger of invention with respect to PhDs is that you may invent something that somebody else has already established, or (even worse) invent something that has long been discredited!

1. The QAA “Code of practice for the assurance of academic quality and standards in higher education”

## What is thesis?

- ➊ A thesis is a series of related arguments, building one on top of another, that seek to establish a proposition (*research question or statement*) in such a way that the proposition cannot be (easily) refuted.
- ➋ The subject of your PhD should not simply be an application (although this may be the evaluation medium).
- ➌ Generic research statements:
  - X is good for Y (in the context of Z).
  - X can be extended to achieve Y (in the context of Z).
  - The application of X to Y produces an improved result (in the context of Z).
  - The adoption of X facilitates Y (in the context of Z).
  - An X approach to the problem of Y avoids the need for Z.

## Example Research Statements (1)

- ➊ The use of variable fan-out index trees coupled with sequential data mining for mobile computing offers efficiency benefits. (X and Y is good for Z)
- ➋ The use of wrappers facilitates multi agent data mining in the context of ease of use and overall effectiveness. (X is good for Y in the context of Z)
- ➌ XML interfaces can enhance the operation of remote microscopy. (X is good for Y)
- ➍ Sub-graph mining techniques can be extended to support image classification with beneficial results. (X is good for Y in the context of Z)

## Example Research Statements (2)

- ➊ An argumentation framework that avoids the need for a supporting knowledge base can have significant operational advantages. (An X that avoids Y is good for Z)
- ➋ The representation of functional RMI scans as quad trees supports effective classification of such scans using graph mining techniques. (The use of X for Y is good for Z)
- ➌ Simple language independent statistical techniques for text mining can produce acceptable accuracy, thus avoiding the need for costly language specific approaches. (The use of X for Y avoids Z)

## DECIDING ON A SUITABLE TOPIC?

- ➊ There are some obvious (possibly related) general reasons for choosing an area in which you wish to conduct your research:
  - Interest
  - Experience
  - Knowledge
- ➋ Identifying a suitable topic within that area is harder.
  - It has to be at the “cutting edge” of research (you can not do something that has already been done).

## Where is The Cutting Edge of Research in Your Chosen Field?

- ➊ Knowing what is at the cutting edge of research in your area of interest is difficult.
- ➋ It requires an in depth knowledge of the area, knowledge that the new PhDer often does not have.
- ➌ In fact the new PhD student often has no idea.
- ➍ Supervisor can help or may suggest a research question.
- ➎ Often takes a good year into the research to identify a research question (so don't worry).

## Pushing at The Knowledge Envelope?

- A PhD should contribute/extend our current knowledge and understanding.
- This contribution/extension thus not necessarily have to be substantial (although of course every PhD student would like it to be).



## CONDUCTING YOUR RESEARCH?

- Research is a slow step by step process of building on top of what has already gone before. (We do not want to reinvent the wheel!!)
- This implies the methodology in our definition.
- Once you have a feel for your research question start thinking about how you are going to work towards clarifying your thesis and then the methodology to establish it.

## A Generic Methodology

Work Package	Duration	Deliverables
Playing around with idea and confirmation of thesis	6	Paper and pencil!" exercises, clear idea of what you are trying to achieve, detailed
Generation of test bed framework	2	usually well engineered software
Experimentation and evaluation	8	Questionnaire returns, Graphs, tables, etc; a deep insight into
Write up (revisit literature review, carry out any additional work)	8	final dissertation

## Background Reading

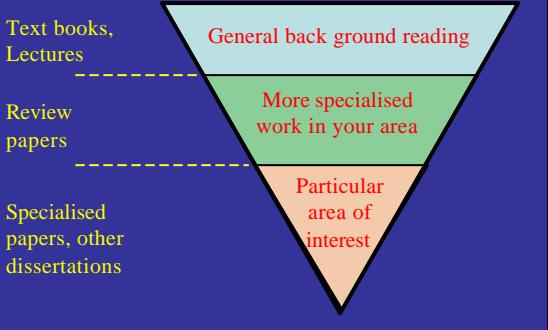
- Most PhD programmes of work start with a “literature” or “state of the art” review.
- Objectives of this are:
  - To be aware of existing work (i.e. not to “reinvent the wheel”).
  - Assist in determining your thesis/research question (try to identify deficiencies in current work, open questions, things that have not been addressed)
  - Become aware of other (parallel) work in the field.
  - At the end of the PhD you should be able to consider yourself to be a (world) expert in your field of study --- thus you need to know your literature.



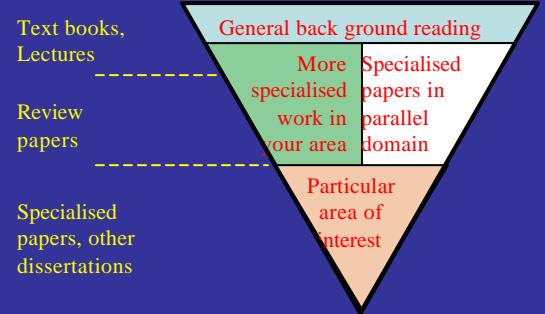
## Knowing your research domain

- Identify the key forums for publications in your domain; there are lots of specialised journals, conferences and workshops etc.
- References to recognised conferences and journals carry more weight than obscure workshops.
- Identify the key players in your field, who are the big names? (Your supervisor might be one!)
- Identify the major research groups working in your field (you may be in one).
- Know your research community (networking).

## Scheme for Background Reading (1)



## Scheme for Background Reading (2)



## Notes on Reading Papers

- ➊ Reading articles in your area of interest can be hard going (especially if English is not your first language).
- ➋ Start simply with text books and review papers --- take care with WWW sites (e.g. Wikipedia) whose provenance may be questionable.
- ➌ Make sure you mark papers you “print off” with all the necessary reference material.
- ➍ Make notes (in computer readable form) as you will need this for your literature survey.
- ➎ Mix reading of papers with other work.

## Evaluation is V. Important

- ➊ The evaluation of your thesis is the most important part of the process as this is where you establish that your thesis is correct!

**Do not neglect it!**

## Notes on Software

- ➊ For many PhD students their research involves producing the largest software system that they have ever written.
- ➋ Use sound software engineering techniques (all the stuff you learnt on your undergraduate study was not just to get your first degree).
- ➌ Commenting your code is especially important. (Functions that seem obvious at the time of writing may not seem obvious 18 months later).
- ➍ Make sure you test your code appropriately (discovering a logic error late in your PhD may invalidate large parts of your evaluation).

## Notes on Experiments

- ➊ Some PhDs can involve a considerable amount of evaluation conducted using experiments.
- ➋ When conducting experiments make sure you have some structured way of recording the results.
- ➌ Make notes on each experiment undertaken (in such a way that when you come back to your notes many months later you can actually understand what the experiment was about). Include:
  - Parameters of the experiment (and why selected).
  - The outcome.
  - The conclusion/interpretation of the results.

## Project WWW site (1)

- ➊ It is a good idea to maintain a project WWW site.
- ➋ Not only does this increase the profile of your work, but also your own profile in the international research community (with respect to your particular research area).
- ➌ The international impact of your work is a factor in the examination of your PhD.
- ➍ Having other people reference your work is an indicator of your international standing (you can help by making it easy for people to find and reference your work).
- ➎ Remember that people who come across one of your papers, if they are interested, will “Google” you to try and find more --- so have a “good” WWW page.

## Project WWW site (2)

- ➊ Your site can include things like:

- Mini biography.
- Research interests.
- Review of your work (including preliminary findings).
- Links to sites and conferences related to your area of study.
- Details of conferences you have attended.
- Links to papers, published and unpublished (refer to unpublished papers as technical reports).
- Software for people to download.
- Data sets (provided you are not infringing any copyright or privacy concerns).
- Anything else you can think of to increase your profile.

## You and Your supervisor (1)

- ➊ Quoting from [1]:
  - “Institutions will appoint supervisors who have the appropriate skills and subject knowledge to support, encourage and monitor research students effectively.”
  - “Each research student will have a minimum of one main supervisor. He or she will normally be part of a supervisory team. There must always be one clearly identified point of contact for the student.”
- ➋ The relationship between you and your supervisor is important --- so work at it.

1. The QAA “Code of practice for the assurance of academic quality and standards in higher education”

## Monitoring Your Progress

- ➊ Universities are required to “put in place and bring to the attention of students and relevant staff clearly defined mechanisms for monitoring and supporting student progress”. (It was not always like this.)
- ➋ But you should still draw up a schedule for yourself.
- ➌ Try and develop a plan (or at least a six month “look ahead”).
- ➍ Monitor your progress, set yourself goals and targets.
- ➎ Keep in regular contact with your supervisor.

## Stay Focussed

- ➊ There is a danger that, as your research progresses, you get side tracked (usually simply out of interest in some other area of work that you come across).
- ➋ Do not allow yourself to get side tracked to the extent that you spend months working on some topic that has nothing to do with your PhD.

## Working hours

- ➊ Draw up a sensible schedule.
- ➋ Working long into the night on a regular basis is bad for your health (and your social life).
- ➌ Try and allow your self weekends off.
- ➍ Take advantage of the social and recreational opportunities that universities offer.
- ➎ If you are from overseas it can get lonely ☺.



## Keeping a Log Book

- It is a good idea to keep some sort of log/lab book to note progress, significant results, goals and objectives for the future, etc.
- Some universities insist that you keep a log book.
- Universities are required to “*provide opportunities for research students to maintain a record of personal progress, which includes reference to the development of research and other skills*”.

## Selecting an External (1)

- It is not up to you to select an external (this is one of your supervisors responsibilities).
- However there is nothing wrong with suggesting people.
- Remember that most universities insist that external examiners are:
  - 1) of senior lecturer status or above,
  - 2) and have some authority in your field.

## Selecting an External (2)

- Take care when considering potential externals:
  1. From overseas, or
  2. From industry;

as they may not be familiar with the UK system.



## Professional Development

- Universities are required to “*provide research students with appropriate opportunities for personal and professional development*”.
- Typically this includes short courses (sometimes only one day) on things like doing presentations, using library facilities, software, etc.

## You are a Scientist!

- Remember that (whether you like it or not) you are a scientist --- a Computer Scientist.
- So behave like a scientist --- be objective at all times.



## WRITING UP

- The principal result of your three/four years of work is your thesis (this is your legacy).
- The thesis (as noted earlier) seeks to establish your research question (statement) and this question should run like a thread through your write up linking the introduction to the conclusions via the intervening chapters

## Generic Format (1)

1. Introduction.
2. Previous work.
3. Details of the framework developed to establish the research question.
4. Details of Experiments and evaluation.
5. Conclusions.
6. Summary.
7. Appendices.

## Generic Format (2)

1. Introduction.
2. Previous work.
3. Application area.
4. Details of the framework developed to establish the research question.
5. Details of Experiments and evaluation.
6. Details of Further Experiments and evaluation.
7. Conclusions.
8. Summary.
9. Appendices.

## Some General Notes

- Write in a scientific manner.
- Use what ever techniques are best to get your ideas across. Do not be afraid of using lists, tables, bullet points, diagrams etc.
- If including algorithms present them formally accompanied by a textual explanation.
- Do not overuse acronyms and symbols as the reader will not be able to remember them all.
- Keep to the point!

## Chapter 1: Introduction

- Introduce thesis and research question (this will require some background work, but leave most of this for the previous work chapter).
- Use the introduction to define the scope of your thesis, what you have and have not looked at and why (you cannot do everything).
- Include your research methodology (how you set about conducting your research).
- Define the criteria for success, how you will demonstrate (prove) your research proposition.

## Chapter 2: Previous Work (1)

- Base your previous work chapter on your “state of the art review” conducted earlier in your PhD.
- Make sure you have kept up to date. (Reference your external examiner’s work if appropriate!)
- Divide the previous work conducted in your field of study into a hierarchy to identify groupings and sub-groupings of work.
- In individual groupings and sub-groupings order developments in chronological order.

## Chapter 2: Previous Work (2)

- Throughout maintain a link with your original research objectives.
- There is a temptation to write down everything you know or have read about your research domain --- stay focussed.

## Chapter 3: Framework

- In Computer Science PhDs the frame work which you are going to used to establish your research question often involves the definition of algorithms, pieces of code or substantial software systems.
- Use appropriate specification and design techniques.



## Chapter 4: Evaluation

- As noted earlier the evaluation chapter is the most important chapter in the entire thesis.
- It is where you establish your research question.
- Evaluation can take a number of forms, but common approaches are:
  - Mathematical proof
  - Comparison with alternatives (often those proposed by other researchers)
  - User feedback in the form of questionnaires
  - Combinations of the above.

## Chapter 5: Conclusions

- Summary (say what you have said).
- Main findings (linked back to the criteria for success itemized in the introduction).
- Suggested for future work.



## Including Code?

- If a substantial software system has been produced it is nice to give the examiner an opportunity to run it. Consider:
  - Including code on a disk, or
  - Place the code on your WWW site and reference this (see earlier discussion).



## Referencing

- Some Universities insist on a specific referencing system (e.g. Harvard).
- If not adopt a system (but remember to be consistent throughout).
- Referencing WWW pages is OK, but take care --- most WWW sites are not refereed (unlike journal and conference papers).
- Remember that (as a rough guide) references to recognised conferences and journals carry more weight than obscure workshops.

## ARE PUBLICATIONS IMPORTANT?

- If you manage to get some publications describing your work this is good --- it means that at least some aspect of your work has been independently refereed and thought to be worthy of publication.
- However publications are not everything.
- Some students get very obsessive about getting publications, sometimes to the detriment of their PhD.

### Are Publications Important? (cont)

- The student at the desk next to yours will always have more publications than you (accept it).
- In your rush to publish be careful who you publish with (avoid “vanity publishing”).
- If in doubt ask around (“impact factors” are a good guide, also “acceptance rates”).
- Normally you work together on papers with your supervisor (or supervisory team).
- Never put your supervisor’s (or anybody else’s) name on a paper without asking them.

## TROUBLE SHOOTING

- There are various things that can go wrong during your PhD.
- Examples:
  1. Funding problems.
  2. Discovering (late on) that somebody else has done the same as you.
  3. Breakdown in Student-Supervisor relationship.
  4. “Brick Walls”.
  5. Departmental problems and issues.
  6. Moments of darkness and despair.

## Trouble Shooting: What To Do.

- It is in nobodies interest for you not to get your PhD.
- Tell somebody as soon as possible (“a problem shared ...”).
- This person should normally be your PhD supervisor (unless your supervisor is the problem).

## Funding Crises

- ➊ Many students have funding problems. If your bursary/grant covers both your PhD funds and all your living expenses then you are lucky.
- ➋ There are various funding bodies all with various strings attached: i.e. only for:
  - Particular subjects
  - A specific university,
  - Students from a particular country or geographic location.
- ➌ Overseas Research Students Awards Scheme (ORSAS), operated by individual universities.

## Competitors

- ➊ If you discover that somebody else has done some similar work to you but you did not know about it till the writing up stage of your work then:
  1. If it is high profile and
    - The work is very recent then report it as a parallel study (nothing wrong with this).
    - Otherwise (not recent) you should have known about it, reference it and state that you were unaware of it at the time of conducting your work.
  2. If it not high profile, i.e. published at some obscure workshop, ignore it.

## Brick Walls

- ➊ In (nearly) every PhD a point is arrived at where the student discovers that he/she cannot progress the research any further because it *appears* (for example) that:
  1. It simply cannot be done (this is why nobody else has done it).
  2. Your software skills are not up to the job (and you suspect nobody else's are either).
  3. You are clearly not intelligent enough and should have gone into refuse collection instead (no disrespect to refuse collection personnel intended).
  4. You have "lost the plot" and no longer know what you are doing.
- ➋ i.e. you have hit a "brick wall".

## Brick Walls --- getting round them

- ➊ When hitting what seems like an insurmountable problem the first thing to do is step back and take a break for a few days.
- ➋ A fresh look after a few days away is always beneficial.
- ➌ Try and get round the problem somehow (i.e. go round the side of the "brick wall").
- ➍ Do not be afraid of backtracking, even if you have spent several months pursuing a particular route of enquiry.
- ➎ You can include negative results in your thesis (but if it is all negative, i.e. the answer to your research question is "no", then think about changing your research question).

## Supervisor-Student Relationship Breakdown (1)

- ➊ It does not often happen, but sometimes the relationship between a student and supervisor breaks down.
- ➋ For example because of a personality clash or academic disagreement.
- ➌ Who do you turn to?
- ➍ Most universities insist on a second supervisor and/or an advisor. There may also be a director of Post Graduate Research (PGR), or ultimately you Head of Department.

## Supervisor-Student Relationship Breakdown (2)

- ➊ Sometime the student-supervisor relationship breaks down because your supervisor has accepted a post at some other institution and you do not want to move to that institution.
- ➋ At least in this case the breakdown is not an acrimonious one.
- ➌ The department does have a certain responsibility to you, and should make appropriate arrangements.

## Departmental Problems

- Sometimes difficulties are encountered with things like: library resources, hardware, out of hours access to buildings, print quotas, etc.
- There should be some formal mechanism for you to raise concerns (e.g. a Postgraduate Staff Student liaison Committee).
- Details are normally to be found in a “post graduate student handbook” that it is now customary for universities to provide.

## Moments of Darkness and Despair (The Mid-Term Blues)

- Every PhD students has moments (days, weeks, ...) of self doubt: “Why am I doing this?”, “What’s the point?”, “Am I wasting three years of my life?”.
- This is a well recognised (almost) medical condition known as the *mid term blues*.
- Mid term* because it typically strikes in the middle of your second year when you can no longer see the beginning, and the end is no where in sight (often coinciding with the brick wall syndrome).
- Knowing in advance that this is going to happen helps to get over it.

## Complaints

- Hopefully things will never get this bad!
- However there is a legal requirement for universities to ensure that: “*independent and formal procedures will exist to resolve effectively complaints from research students about the quality of the institution’s learning and support provision*”.
- Also to provide “*formal procedures to deal with any appeals made by research students*”.
- Again details are normally to be found in your PG student handbook.

## SUMMARY

1. What is a PhD?
2. Deciding on a suitable topic
3. Conducting your research (identifying a methodology)
4. Monitoring progress (goals and targets)
5. Writing up
6. Are publications important?
7. Trouble Shooting
8. Summary and conclusions

## CONCLUSIONS

- Doing a PhD is terrific fun.
- It is a once in a life time opportunity to apply your intellect and wit against a significant challenge.
- So relish it.

