Hints and Tips on (Science and Engineering) Bachelor's and Master's Thesis Writing

Peter McMahon[1]

May 2009

1 Introduction

As the title suggests, this document isn't a comprehensive guide to thesis writing. Instead it is a collection of random pieces of advice that I would have liked to have been told before I began writing my first thesis. I think most of the advice is applicable to students writing theses in all fields of quantitative sciences and engineering. However, none of what follows has been rigorously tested — everything presented is my uninformed opinion and I'm sure would have benefitted me, but perhaps may be less helpful to others. *Caveat emptor*, and good luck!

2 Strategies for Thesis Writing

Students who begin writing their first thesis often make a critical mistake: they start putting down words immediately, without first planning their thesis. You should plan first. Make a high-level outline. Once this is done, go through each chapter and add subsections. I recommend that you then go through each chapter and add your figures or placeholders for figures. I used this technique when I wrote my Master's theses, and I found that it made writing much easier: the bulk of the writing I needed to do was just that for describing each figure. By deciding first on what figures you will include, you will have compiled a list of figures that you still need to create. If you begin by preparing most of your figures, you'll reduce your chances of being surprised by missing data later on.

Write your "background" chapter last. The reason for this is quite simple: it's very easy to spend an enormous amount of time writing this chapter, and leaving it until you're done with everything else makes sure that if you spend too much time on it, the presentation of your actual work won't suffer consequently. It's very tempting to make your review of background material a comprehensive tutorial of your field, because it's fun to write and you get to look up and learn new and interesting things as you go

[1]UCT ECE B.Sc. 2006; EE, CS M.Sc. 2008. Please feel free to send comments and corrections to pmcmahon@stanford.edu.

along. In contrast, presenting your results is much more difficult: typically by the time you begin writing up, you know so much more about your work than the average reader of your thesis will, that it's a daunting prospect considering how best to coherently present everything you've done in the best light possible. Don't use your background chapter to delay writing the meat of your thesis!

Write! At some point, you have to stop doing experiments. It's tempting to keep working on your experiments because there are always extra things you can try. However, at some point (typically at least one month before your thesis handin deadline) you have to transition from doing experiments and analysis to writing up. As you write, there will be times when you need to "quickly" do some extra work when preparing a figure, or a table of results, but you should resist the temptation to do anything more that isn't absolutely necessary. Your readers would much rather read a well-written account of your $N-4$ weeks of research in your thesis than a rushed and sloppy account of $N-2$ weeks of work (where N is the total duration of your thesis in weeks). Your thesis report is likely the only means your reader will have to evaluate your work, so make sure your writeup reflects the same dedication to quality that you put into your research!

Until you start writing up, you won't know what data or knowledge you're missing, so starting your writeup early will help you avoid a rush before your deadline. It's very annoying discovering a week before you need to hand in that you forgot to run an experiment with some set of parameters, and you don't have enough time to do it with just a few days left. At best, you may have to submit your thesis with some graphs missing data points, and at worst the missing data points would have completely changed your conclusions. Either way, this is undesirable and can be avoided if you find out well in advance of your handin date that you need more data.

3 The Structure and Style of Theses

The best way to learn about appropriate style for thesis writing is to read some theses. (See the next section for suggested reading.) Academic style is quite different from all other forms of writing, even report writing, that you've likely seen before starting your research. It takes a while to get used to. During your thesis, you should have read many academic papers from journals and conferences — these are typically written in a style very similar to that which is appropriate for theses.

There is some room for creativity in structuring one's thesis, but good theses usually have the same basic divisions and flow:

- **Introduction** – $\sim$10 pages. Explain the purpose of the thesis.

- **Background** (Related Work) – ∼20 pages. Provide a summary of the subfield you're working in, to provide some context for the reader to put your work in. Discuss related work done by other researchers, and explain how yours relates to it.
- **Experimental Design** – 20-30 pages. Explain the specific problems you set out to solve, and describe the design of your experiments.
- **Results and Analysis** – 20-30 pages. Present your results, and provide analysis of them.
- **Conclusion** – ∼5 pages. Summarize what you did, what worked and what didn't. Discuss potential future work that may fill gaps in your work, or approaches that seem promising to overcome problems you encountered but that you weren't able to tackle.
- **Appendices** – Include any material that would break the flow of the main body of your thesis in appendices. For example, appendices often include presentation of non-essential data; long proofs; elaborations of ideas for future work; code listings or other "methods" information (circuit designs and suchlike), and other documentation such as user guides.

The structure and style of a good thesis typically looks very different to what is taught in "professional communication" classes in many universities. This is unsurprising: in most cases, professional communication lecturers are neither scientists nor engineers, and have never written a scientific document, let alone a science or engineering thesis. No wonder their advice is ill-suited for science or engineering writing! Any grammar lessons that one obtains in professional communications classes are advantageous, but I recommend that you ignore everything else that they teach about report writing.

If your thesis is mathematical, then I can recommend a guide to mathematical writing co-authored by Donald Knuth. Stanford offered a course in 1987, "Mathematical Writing", that was co-taught by Prof. Knuth, and a transcript of the course was released as a technical report that makes for very pleasant reading. There are many excellent 'before' and 'after' examples; the course is packed with practical advice to make your writing easier to understand.
URL: ftp://reports.stanford.edu/pub/cstr/reports/cs/tr/88/1193/CS-TR-88-1193.pdf

4 Examples of Outstanding Theses

- University of California, Berkeley David J. Sakrison Memorial Prize for a Ph.D. Thesis (2005): "Limits on Efficient Computation in the Physical World" (Scott Aaronson). This is the gold standard by which all theory theses should be measured — it is a very fun read, contains excellent overviews of background material,

and clearly explains all the results that were obtained.
URL: http://arxiv.org/abs/quant-ph/0412143

- MIT Ph.D. Thesis (2001): "Photonic Crystals: From Theory to Practice" (Steven G. Johnson[2]). This thesis is the gold standard for all physical sciences and engineering theses. It is witty, but gives the reader a very serious and intuitive introduction to the subject matter, placing the achieved results in context in a very effective manner.
URL: http://hdl.handle.net/1721.1/8644

Meeting the high standards set by Aaronson's and Johnson's theses is not a realistic goal for most of us (although it is good to know what superstars of thesis writing can achieve). The following theses are also examples of great style and structure, and are somewhat more conventional than the gold standards. I have tried to pick out theses from several different areas.

- University of California, Berkeley M.S. Thesis (2008): "RDLC2: The RAMP Model, Compiler & Description Language" (Greg Gibeling). This is an outstanding Master's thesis that provides a very clear and comprehensive coverage of the material.
URL: http://bwrc.eecs.berkeley.edu/php/pubs/pubs.php/39.html

- Association for Computing Machinery Doctoral Dissertation Award Honorable Mention (2008): "Seeing the World Behind the Image: Spatial Layout for 3D Scene Understanding" (Derek Hoiem).
URL: http://www.ri.cmu.edu/publication_view.html?pub_id=5825

- Association for Computing Machinery Doctoral Dissertation Award Honorable Mention (2007): "Conditional Graphical Models for Protein Structure Prediction" (Yan Liu).
URL: http://www.lti.cs.cmu.edu/Research/Thesis/YanLiu06.pdf

- MIT David Adler Memorial Thesis Prize for Best Electrical Engineering M.Eng. Thesis (2007): "Capacitive Sensing with a Fluorescent Lamp" (John J. Cooley).
URL: http://hdl.handle.net/1721.1/41255

- American Physical Society Atomic, Molecular or Optical Physics Dissertation Award (2005): "Ultracold Bosonic Atoms in Optical Lattices" (Ana Maria Rey).
URL: http://hdl.handle.net/1903/1802

[2]Besides Prof. Johnson's work in computational methods for investigating photonic crystals, he is also known for creating the Fastest Fourier Transform in the West (FFTW). He has co-authored an outstanding textbook on photonic crystals that is available for free online as an e-book, and is another superb example of technical writing. URL: http://ab-initio.mit.edu/book/photonic-crystals-book.pdf.

- MIT George M. Sprowls Award for Outstanding Ph.D. Thesis (1999): "Mostly-Static Decentralized Information Flow Control" (Andrew Myers).
  URL: http://hdl.handle.net/1721.1/16717

- MIT William A. Martin Memorial Thesis Award for an Outstanding Master's Thesis in Computer Science (2006): "Probabilistic framework for genome-wide phylogeny and ortholog determination" (Matthew D. Rasmussen).
  URL: http://hdl.handle.net/1721.1/37921

- MIT William A. Martin Memorial Thesis Award for an Outstanding Master's Thesis in Computer Science (2005): "End-user programming for the Web" (Michael Bolin).
  URL: http://hdl.handle.net/1721.1/33110

- MIT William A. Martin Memorial Thesis Award for an Outstanding Master's Thesis in Computer Science (2004): "Tree Pattern Inference and Matching for Wrapper Induction on the World Wide Web" (Andrew W. Hogue).
  URL: http://hdl.handle.net/1721.1/28406

- MIT William A. Martin Memorial Thesis Award for an Outstanding Master's Thesis in Computer Science (2001): "Adaptive Delivery of Real-Time Streaming Video" (Nicholas Feamster).
  URL: http://hdl.handle.net/1721.1/8570

- MIT M.S. Thesis (1993). "Robust, High-Speed Network Design for Large-Scale Multiprocessing" (Andre DeHon).
  URL: http://hdl.handle.net/1721.1/6790

- MIT Ph.D. Thesis (1996). "Reconfigurable Architectures for General-Purpose Computing" (Andre DeHon).
  URL: http://hdl.handle.net/1721.1/10637

- University of California, Berkeley Ph.D. Thesis (2001): "Decoherence, Control, and Symmetry in Quantum Computers" (Dave M. Bacon). This is another fun thesis that demonstrates the possibility of presenting difficult, interesting material in an engaging and entertaining fashion.
  URL: http://arxiv.org/abs/quant-ph/0305025

Many professional societies have thesis awards, and these can be a useful source for finding other theses that have been recognized as excellent. For example, the American Physical Society has AMO awards, and the Association for Computing Machinery has a Doctoral Dissertation Award. The ACM award is given out annually to what is deemed the best Ph.D. thesis in the world in Computer Science submitted in the past year[3].

[3]Most often the winners are from MIT, UC Berkeley, Stanford, Carnegie Mellon, and the University

Theses need not have boring layout. If you are adventurous and would like some inspiration for unconventional layout, you can refer to the undisputed leader in thesis presentation:

- Utrecht Ph.D. Thesis (2004): "Radio Pulsars"[4] (Joeri van Leeuwen).
  URL: http://igitur-archive.library.uu.nl/dissertations/2004-0512-101234/full.pdf

5 Technical Tools

Take the time to learn LaTeX — you won't regret it! If you're not very technically inclined, then you might like a WYSIWYG LaTeX editor, such as lyx. However, you get significant speed advantages from writing directly in LaTeX markup. Another major advantage of writing raw LaTeX is that you can focus on the content of your writing and worry about the layout later. When you see how your material is being presented in a WYSIWYG editor, it is very tempting to continually tweak the layout, which distracts you. With LaTeX, you can write your text and then only worry about making the layout look just the way you want it to.

Cambridge University's Engineering Department has a comprehensive collection of LaTeX resources. I recommend you use this as your starting point; it has links to both Windows and Linux installation/usage instructions and both beginner and advanced LaTeX tutorials. URL: http://www.eng.cam.ac.uk/help/tpl/textprocessing/

Good figures are critical to a thesis. If you use unwieldy tools, you'll become frustrated and undermine your work by producing subpar figures. The following software is useful for producing good-looking diagrams for theses:

- OmniGraffle (Mac OS; commercial) — hands down the best diagramming software available, but unfortunately restricted to Mac OS.
- Microsoft Office Visio (Windows; commercial) — the best option for Windows users who have free access to it. An unconvincing choice if you have to pay the retail price.
- Microsoft Office PowerPoint (Windows; commercial) — can be used to make appealing block diagrams, but is found wanting for anything more complicated.

of Washington, but occasionally theses from outside of the United States are recognized. The Technion, and Weizmann Institute in Israel, and Ecole Polytechnique Federale de Lausanne in Switzerland have produced winners or runners up in recent years.

[4]This thesis also includes a very accessible introduction to the field of radio pulsars, although it is unfortunately only available in Dutch. The main body of the thesis presents, in English, a very readable summary of each of the papers the author wrote during his studies.

- InkScape (Windows/Linux; free) — not bad, but has many flaws. Quite possibly the best free diagramming tool.

Make sure you use vector image formats for your diagrams and graphs. For figures that are intrinsically in raster format (such as photographs of experimental setups, or image outputs of computer simulations), use the highest resolution images possible — few things create a worse first impression for a thesis reader than the appearance of pixelated images every few pages.