Design:

The Key to Writing

(and Advising)

a One-Draft Thesis

John V. Carlis

Computer Science and Engineering

University of Minnesota

{carlis@cs.umn.edu}

September, 2002

Introduction.

If you are a Ph.D student who needs to write a thesis proposal or a thesis itself then this paper ought to interest you, and your advisor too. The advice here also applies to you if you merely (hah) need to write a research paper.

My goal in this paper is to codify an effective and efficient way of work, so that I more readily can pass it on to my own students, and can share it with other students and their advisors.[1]

What's my story? While working on my thesis, and at the same time a professor at Syracuse, several streams of thought happily converged. I read several books on writing. I read the forward to the John McPhee Reader, which described his disciplined, design-before-write way of work[2]. I read Richard Mager's Preparing Instructional Objectives, which has this message: teaching (writing) is about them, the students (readers), not me. I taught software development using Yourdon and Constantine's "Structured Development" (Addison-Wesley, 1975). Convergence began when the strong parallels in their content struck me.

I already believed that software should be designed and not hacked, and came to the same conclusion about a thesis. I decided to try to transfer software notions to writing. Since a paragraph is a unit of development, like a software procedure, I chose to design down to the paragraph topic sentence level before writing. And it worked! Now, as an advisor, I coach my students so that they too can successfully design and write a one-draft thesis.

---

[1] I was telling some "Bushies", mentees of mine in the University of Minnesota's Bush Early Career Faculty Development Program, about my one-draft notions when Patricia Schaber said "Come talk to my department about it". That invitation motivated my writing this paper. So, thanks Patricia.

[2] McPhee won the Pulitzer Prize for "Annals of a Former World", Farrar Straus & Giroux, 1998. Besides writing great books, he teaches writing at Princeton. Read his books.

What makes me qualified to give advice about a one-draft thesis? Experience does. I partially did it. Four of my advisees (Leone Barnett, Libby Shoop, Paul Wagner, and Yue (Jake) Chen) have done it.

What can motivate you to follow my advice? Four things can. First, since a thesis is territory unknown to you, you should be wary, if not scared, and eager for advice on how to succeed. Your thesis most likely is, as it is for most students, the largest writing project you will ever undertake. Unless you are one those rare people who already has a Ph.D., I doubt that you have the experience to say "no problem". Second, you may find writing painful; many students do. Sadly, after years of hard work, they struggle to produce a barely-acceptable thesis. So you ought to be willing to listen if doing so will ease your pain. Third, *ceteris paribus*, your committee will consider your thesis as having higher quality substance if it is better written. So you will benefit by writing well. Fourth, the thesis is more important than other writing that you have done — your career hangs in the balance because: no thesis, no faculty job.

Think about this: I claim that if you follow my design way of work, you can produce a one-draft Ph.D. thesis. This happy thought ought to make you willing to read on. Begin by studying the summary on the next page which briefly states what you need to do to obtain that success, and the key notions that I want you to understand and employ. It also delineates the organization of the paper.

Beware — you cannot just casually read this paper, and then, poof, you make a thesis appear. If you are going to succeed you will have to think hard about ideas that may be new to you. You will need to practice. You may need to break some old, bad habits. You may need to become disciplined, because just winging it doesn't work for a thesis. You may need to educate your advisor, which may be your hardest task.

Summary

You succeed by seeing the whole thesis as flesh on the bones of the design, with contributions the *raison d'etre* of the thesis.

To Succeed:		Key Notion
1. Believe	In a one-draft thesis.	Gumption
2. Contribute	To the field.	Destination
3. Avoid	The fatal writing error: wrong audience.	Sketch
4. Acquire	Writing skills.	Paragraph
5. Design	A thesis.	Avoid False Progress

1. Believe   In a one-draft thesis.

In order for you to be willing to try to write a thesis in one draft, you must believe it can be done at all, and, more to the point, that *you* can do it. To acquire that faith consider the following questions.

What does "draft" mean? A draft is a completed work, something that you give to others for review. I do not consider small scale revising, say, editing within a paragraph while leaving the structure the same, as re-drafting. Do not misunderstand; "one draft" does not mean that you just start from a blank slate with final words of truth and beauty rolling off your fingertips. No, producing a thesis is work, but it is merely work, not some mystical thing.

Can it be done? Absolutely, since four of my students did it.

What specifically should you believe? Three things. First, believe that re-drafting is both costly and avoidable. By writing only one draft, you can save at least several months worth of effort. Second, believe that one draft is preferable to several drafts. It indicates good organization and yields a more readable thesis. It does not indicate of lack of effort – beware, you may have to fight your own or your advisor's preconceptions about this point. As you will see, I consider re-drafting to be a consequence of incomplete or flawed design. Third, and most important, believe that indeed *you* can write a one-draft thesis. I hope to convince you that if you can write a paragraph — and surely you can or can learn to do so — then you can write a one-draft thesis.

What benefits will accrue to you if you do believe? It will raise your gumption level[3].
When you understand what is going on and why, then you will have gumption. If not, then
you can worry and perhaps do no writing — those with gumption do not suffer much from
writer's block — or do wasteful things in the vague hope that good things will occur.
Also, the way of work you will learn will help you with other writing tasks, such as writing
research papers and proposals for funding.

---

[3] See R. Pirsig's "Zen and the Art of Motorcycle Maintenance", Morrow, 1974.   It's a great book;
philosophy with a plot and characters.   Read it.

2. Contribute To the field.

You need to understand that

- The purpose of a thesis is to defend your claimed contributions to your field.

In other, more dramatic words:

- No contributions, no thesis!

Where do contributions come from? Your abiding interest in a topic and hard work enable you to have something interesting and verifiable to say to others. To obtain contributions you might state a scientific hypothesis that you test, or prove a theorem, or engineer a solution to a problem, inquire into the state of some phenomena, or use some other scholarly mechanism.

When should you think about contributions? Well, here is a major mistake:

- You first write the bulk of the thesis, and then try to determine what contributions you can claim.

That's the tail wagging the dog, and will cause you great trouble. Instead, you should begin thinking about contributions much earlier than draft-writing time. In your proposal you have target contributions. In pre-proposal pondering you try out possible contributions. You imagine potential "BIG WOW!", that is, if all goes well, what would turn your field on its ear. If you have no chance of a big wow, pause and examine where you are heading. Also, assess risk: if something goes wrong, will you still graduate? In short, you think about potential contributions from the very beginning. As you proceed

you change the phrasing from "what could I perhaps contribute" while musing to "I intend to contribute" in the proposal to "I have contributed" in the thesis.

What constitutes a contribution? Well, in spite of its importance, many students, and some advisors, misunderstand "contribution". To grok[4] it consider these three important terms: novelty, consequence, and evidence. To your committee, who represent the other[5] experts in the field, you will say "Look, here is something new and valuable, that somebody cares about, and I can prove it". Your immediate goal is to get the committee to sign the paperwork allowing you to graduate. You do that by informing them of what your contributions are, and by persuading them that your contributions are valid and significant, and maybe even useful.

Novelty.   To persuade them about novelty, you present a literature review, an organized critique of who has done what, their contributions, and the limitations of their contributions. You need to include the relevant literature, but exclude the irrelevant literature — thereby showing that you are an informed expert who can make judgments. You need to present it cogently, succinctly, pithily, and other –ly words to show that you are not a ninny, and can distinguish dross from diamonds. Since you are exploring unfamiliar territory, you should expect to resolve vocabulary mismatches, e.g., "Smith 1990 coined the term X, but Jones 1991 used X in a subtly different fashion … ". You should expect to invent new vocabulary, and then precisely define it – every thesis makes fine distinctions among terms. You demonstrate expertise by picking an appropriate vocabulary to focus on, and organizing it. The literature review also serves as the substrate upon which you show the goodness of your additions. In your proposal a literature review helps you describe your potential contributions. (More to come on this topic in section 5.)

Consequence.   To persuade your committee about consequence, you state who cares about your novel contributions, and what impact your contributions have on them. Of

---

[4] Grok was coined by Robert Heinlein in "Stranger in a Strange Land", Ace Books, 1961. It means "deeply understand".
[5] By the time you defend your thesis you must be an expert. So act the part.

course *you* care, but that is immaterial. If you produce something new, but it has no impact, it is not valuable to the field, and, therefore, not a contribution.

Avoid the error of thinking that a contribution is about you — it's about them, the people in your field. So do not stop at saying "I did all this work", because the reaction you get is likely to be "So what?" One student/faculty candidate wrote a 50,000 line program for his thesis, certainly a massive effort, but it was more like contract programming than research. So be sure that you remember to distinguish between means and ends, and that it's the ends that count. Also do not stop at saying "I learned this", because *your* learning something is not a contribution to the field if they already know what you just learned — we call that education.

Instead, say things like "My explanatory model for … is better by …", "My new method yields 50% more … than the best previous method", or "My new operators enable users to state …, which, previously, they were unable to state".

Evidence.   To persuade your committee to accept your claims about the impact of your novel contributions, you need to provide evidence. If you invent or discover something new that will, if true, have impact on the field, but you have no evidence, then you have no contribution. Evidence persuades the committee that you have been scholarly, and have fairly assessed the validity and strength of your contributions.

How should you use contributions? From beginning to end, keep the contributions in the forefront of your thinking. Keep a copy in front of your eyes -- literally. Critically review them often. Polish them as you design; it is well worth the effort. Talk to others about them, because, if you can explain your contributions to somebody outside the field, then your committee will be more likely hear a clear explanation from you and then smile on your work.

You should formulate your contributions early, because when you know what contributions you intend to defend you can more easily do these important things:

9

- Figure out when you are done.

  Quite a few students plead "I'm tired. Can I graduate?", or "I've been at this for six years, can't I stop now?", or some such irrelevancy. You are done when you have defended contributions — not before, and, happily, not after.

- Protect yourself from your advisor.

  A student/faculty candidate, interviewing one January, said "I can be here Fall semester. I want to be done in May, but my advisor said, 'No, not until December'. We compromised on late September". Don't let this happen to you — schedule doesn't matter, substance does.

- Decide what and in what depth you must write about, and what you may skip.

  Of course, you will have lots to say, but, if the audience, your committee, already knows and believes something, or it is not relevant to the contributions, then you do not need to produce a treatise on it. I knew a student who added forty pages to his thesis saying "Yeah, I know this is not germane, but I want others to read it, so I put it in anyway". Do you think he wasted his time? I do.

- Decide what other work you need to do and not do.

  Suppose you want to claim a particular contribution, but cannot prove it yet. You must drop the claim, or weaken it, or do more work to gather evidence, or move it to a future work section. Do not spend effort on something just because you could do it. For example, a student did irrelevant work, spending a year devising evidence for his novel software, but, sadly, he picked a sample problem that could not possibly have shown the goodness of an intended contribution. He was poorly advised.

How much is enough? There are no firm, universal rules about just how much of a contribution is needed for a thesis. Disciplines differ, as do advisors. So talk to your whole committee. Do not just propose to take something you have already done (perhaps a master's project), and, without considering the context, say "For my thesis I'm going to add this bit". Don't try to just move out from where you are. That puts the focus on you, and not on the field where it belongs. Also, do not merely propose to make three, small,

weakly or unconnected contributions, and claim that they make a thesis. A thesis is not a grab bag, so avoid this mistake. (In some fields or with some advisors the thesis is just a slapping together of three papers, co-authored with the advisor, and not an integrated work. Be sure that you ask so that you avoid misconceptions.)

Destination. Think of contributions as the *destination*. Research explores new territory (or re-envisions old territory). When you reach the destination you are done. While you still have to choose a path to get there, and that will take work, at least you will not wander. You wander if you work on notions that do not relate to your contributions.

Suppose you are driving, and come to a fork in the road. If you don't know where you are going then you cannot sensibly decide whether to turn left or right. A thesis writer without contributions faces the same dilemma.

Think in terms of destination when somebody says "tell me about your thesis". If you reply with "I'm studying XXX; it's fascinating and I'm learning a lot", then you are thinking as a student, not as a research scholar. You are thinking in terms of your own movement, and not about the territory — and what's new to you may not be new territory to the field. Yes, of course, you indeed will have an abiding interest in the topic. Don't work on it unless you find it interesting enough to continue working on it for years after you finish the thesis. Yes, you will learn, but that's irrelevant for research. If you reply with "I'm hoping to find something new about XXX", then you are thinking about new territory, but only about direction, not how far in that direction. Do reply with "Look, here is something new and valuable about XXX, that these folks in the field care about, and I can (or will) prove it this way".

Use this notion of destination when you are in pre-proposal or proposal stage. One of my students (Paul Wagner) wanted to work on a particular topic but he could not find an appropriate dataset on which to test his notions. Because he knew early that he would not be able to get the needed evidence to support his intended contributions, he wisely picked a different thesis topic, and saved himself a lot of pain.

3. Avoid The fatal writing error: wrong audience. *Sketch*

Not all writing is the same. In particular, writing for yourself qualitatively differs from writing for others. In stick figure form:

                                    is not   .
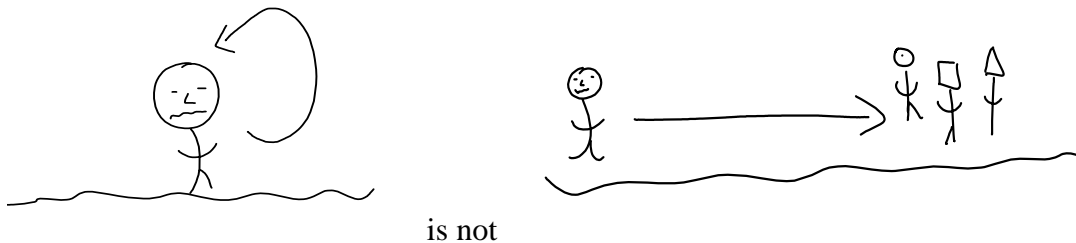
When you write for them (the audience), you put forth effort to make your product good: neat, complete, proffed (sic), unified, cogent, etc. That is, you craft and polish a well-designed product for them to read. However, if *you* are the audience such effort is a waste.

Sketch. To help you avoid the fatal error, think of writing for yourself as *sketching*. If you do, then you expect several things:

- Changing sketches — do not try to get it right the first time.

- Incomplete sketches — do not feel driven to finish them.

- Messy sketches — do not waste time making them pretty.

- Discarded sketches — keep the ideas, not the medium on which you jotted them down.

Why sketch? Well, you can have only a vague idea of your destination (aka contributions), and you sketch to generate ideas. You can sketch to try out a collection of terms, or an exposition strategy to see what its strengths and weaknesses are, so you can compare it to other things you have sketched. Do not make the mistake of thinking that sketching is a sign of weakness – Leonardo Da Vinci sketched. Do not avoid sketching in order to save

13

time -- you will cost yourself time because you will discard what you pretended was a final product.

Blueprint. Sketching is important, but you need more. From sketches you design the thesis — still writing for yourself. Think of it as coming up with a blueprint, which is a more formal design than a sketch, but is still clearly not a house. A blueprint shows the important design decisions that affect the building of the house, and shows enough detail so that competent workers can build the house. A thesis's detailed outline is like a house's blueprint. You design to a level of detail where you, as a competent writer, can confidently proceed to construct the draft.

Why is this audience distinction important? Well, you care about cost, and you can change a sketch or a blueprint or model cheaply (in time, effort, or money), but it will cost you dearly to change the house. Moving a window on a sketch costs pennies, perhaps dollars on a blueprint, but on a house perhaps thousands of dollars. If you try to write without a design you will end up rewriting, tossing away expensive words — so don't do it. Furthermore, when you work from a blueprint, you can concentrate, focusing your attention on building one small piece, and ignoring almost everything else. Likewise, redrafting is expensive, but modifying your design is relatively cheap. Furthermore, the distinction's importance grows nonlinearly with size (a thesis is indeed big), and your inexperience. You would not build a house without a blueprint, so do not write a thesis without a design.

Duh! Nobody confuses the cardboard model or the blueprint of a house with the house. Everybody agrees that it is a waste of time (foolish) to make the model's toilet flush. However, many forget, when writing, that designing is not the same as drafting — because both activities use the same material, namely, words. Because of this error I do not use the term "writing to learn" as Zinsser[6] does, preferring "sketching to learn" instead. My advice then is: don't fool yourself, and commit the fatal error of confusing the blueprint (design) with the house (draft). (Expect more on this topic in a bit.)

---

[6] William Zinsser, "Writing to Learn", Harper & Row, 1988.

But how detailed a design? The next two sections answer that question.

4. Acquire Draft-writing skills.   *Paragraph*

Are you already a good writer? If you are, well that's terrific, but you should expect that writing your thesis will challenge you anyway. If you are not, then you better exert yourself to get better, and do so now. Don't wait until you are partway through a bad draft to begin remedial training. Some students are unaware of how poor their writing skills are, while others think it is no big deal[7] — it is —, and still others think that writing is a genetically endowed trait — it is not. Fortunately, you can acquire the skills that you need.

By the way, do not be smug if English is your native language. I have seen a number of such students who have put their theses at risk due to poor writing skills. On the other hand, a limited idiomatic vocabulary will not prevent you from communicating your contributions — Jake Chen, who was relatively new to English, wrote a fine, one-draft thesis. After all, you don't need to be a Hemingway to write a thesis. You need proficiency in a *craft*.

Paragraph.   The core of that craft is the notion of a paragraph. The good news here, and it is really good news, is: if you can take a paragraph topic sentence, PTS[8], and craft a paragraph of sentences, then you can write a thesis! The potentially-bad news for you is that many students struggle to write a paragraph, and few think explicitly in terms of paragraphs.

You need to be able to:
- Write a good a PTS.

---

[7] A student said to his committee "I have good ideas, I don't need to express them well". Not surprisingly, his committee frowned on his thesis and sent him off to re-write.
[8] Barnett and Shoop both got in the habit of shortening "Here is my detailed design down to paragraph topic sentences" to "Here are my PTSs". So now I use the acronym too.

- Develop one PTS into one paragraph.
- Connect paragraphs together to tell a story.
- Recognize structural errors: poor or missing PTS, and a mismatch between a PTS and its paragraph, especially thoughts that do not support the PTS and belong elsewhere.
- Use paragraph as a unit of thought.

The last bullet is, admittedly, vague, but you can acquire this skill. You can become better at it just by trying, without any particular training. (Expect more on this in the next section.)

Furthermore, learn to be patient in exposing the audience to your ideas. Don't have words gush forth with no designed flow. Become good at using glue words, connectors that allow your ideas to flow to the audience within and between paragraphs. For example, in the paragraph three after this one, look at how I glued things together with controlled repetition of "just".

Of course, you need more than paragraphs, but here, in only a few pages, I cannot even begin to address the language skills that you began to acquire as an infant. However, I can tell you about several prime actions that you can take to improve your mastery of the writing craft.

First, read about writing. Go to the library, and check out a handful from the thousands of books about writing. Read a little bit every day, picking out pearls of advice. I recommend these books[9]:

> Gerald Alred, Charles Brusaw & Walter Oliu, "Handbook of Technical Writing", 6th Edition, St. Martin's Press, 2000.

---

[9] I also recommend these books: a) Joseph Williams, "Style: Ten Lessons in Clarity and Grace", 6th Edition, Addison Wesley, 1999; b) Edward MacNeal, "Mathsemantics: Making Numbers Talk Sense", Penguin, 1994; c) Barbara Tuchman, "Practicing History", Ballantine, 1981 (Tuchman won a Pulitzer Prize for "Guns of August", but suffered terribly in her writing process.) ; and d) D.A. Fryxell, "How to Write Fast (While Writing Well)", Writer's Digest Books, 1992. (He likes McPhee's way of work too.)

Robert Ian Scott, "The Specific Writer", Intl Society for General Semantics, 1995.

Rudolph Flesch, "The Art of Plain Talk", Collier Books, New York, 1951.

Second, learn larger scale templates to mimic. To avoid spewing words, you need to own tools, organizing principles, so that you can craft paragraphs, and sections and chapters too. The books by Alred et al, by Scott, and by Flesch provide some valuable ones. Alred, et al present a number of methods of development: order of importance, cause and effect, sequential, chronological, increasing importance, decreasing importance, division and classification, comparison, spatial, specific to general, and general to specific. They present four forms of discourse: exposition, description, persuasion, and narration. Scott suggests that you see writing as connected sets of questions and answers. He also has chapters about organizing: within paragraphs, sequences, comparisons, arguments, inferences, and figures of speech. Flesch invented a "readability formula" that some word processors implement. It's a rough measure, but with it you can gauge how difficult others will find your writing. He also explains live, crowded, and empty words. Get these books, put their notions in your writing craft toolbox, and use them.

Third, practice smaller-scale basics. Whatever your current level of skill, act like a baseball player going to spring training. Review fundamentals until you are sharp and smooth. Use repetition to inculcate those fundamentals. Slow down and look closely at small things. Become sensitive to the nuances of language, selecting exactly the right words to express precisely what you intend. Doing so can make you aware of how imprecise your current phrasing is. (It still happens to me all the time.)

Of course you must have command over sentences, including even smaller scale basics: verb, voice, and person. Write sentences with powerful verbs in active voice. Don't turn a verb into a noun and then accompany it with a weak verb. For example, don't replace "Julia transferred simply" with "Julia's transference had no complications". Avoid passive voice, because with it you more likely will produce vague statements. Use active voice instead. (You may have to fight with your committee on this one. So prepare yourself to use this argument: passive voice is not more authoritative, it just makes readers work

harder.) Don't be afraid of writing in first person; "I" is not forbidden. I use "I" when I need to, but that seldom happens because I write for "them", my audience. (So, dissect — see below for more on dissect — this paper focusing on person.)

Try this: take some finished writing of yours and highlight the verbs. If you see lots of "is" or other forms of "to be", or "has", then others will perceive your writing as weak.

To help you avoid muddy writing, pay attention to these pairs of "level" words: type or instance; singular or plural; active or passive; set or sequence; and general or specific. Level distinctions form a significant share of the skills I teach in my books.[10] The books are about computer systems people learning how to precisely communicate with users about what they want to remember in a database and what analyses they want to perform. The books are valuable because the limiting factor in database work is not technology, it is communication. I think that good level skills will help you with writing your thesis.

Fourth, just do it, pretending that you are indeed writing for the people in your field. Don't just sit and worry. Practice to get better. Don't just sit waiting for the muse to strike. Work, because thesis writing is craft, not art. Don't just sit idly, waiting to begin some massive final push. Practice writing daily just as you would if training for a marathon. Don't just write something once and then move on to something else. Write the same content several different ways and then critique their strengths and weaknesses. Write the same content for different audiences and likewise critique. Don't just write words, read them aloud to yourself to see if they flow alright.

Fifth, read good writing, and dissect it. You can become a better writer by looking closely at the structure of something you enjoyed reading. For example, read Craig Partridge's excellent book "Gigabit Networking" (Addison-Wesley, 1993). While its content interests me and my students, you can focus on its design, selecting a chapter and reading just the paragraph topic sentences (each begins a paragraph). Those PTSs tell the chapter's story,

---

[10] John Carlis and Joseph Maguire, "Mastering Data Modeling", Addison-Wesley, 2001.
  John Carlis with Scott Krieger, Mastering Database Analysis", Addison-Wesley, 2003 (in press).

but, of course, in less detail. Each paragraph fleshes out the topic sentence, and nothing else.

Read a bunch of theses. Find well-written ones, not considering their relevance to your field. Dissect the writing, and mimic relevant structures. Find poorly written theses, dissect them, and avoid the errors you discover. Find theses in your field, dissecting to get a sense for what kind of things committees have accepted. However, do not feel bound by what others have done.

While you cannot determine the writer's way of work that produced a well-written piece by dissecting it, there's hope — read the next section.

5. Design   A thesis.   *Avoid False Progress*

Here are the parts of my "Design a Thesis" way of work.

- Remember the target audience (you or them), while you write three products (contributions, design, draft).

- Work backwards from contributions to the field, not forward from what you know or what you did.

- Choose vocabulary, both technical content and glue.

- Use figures both to find and show organizing principles, and to present details.

- Design a story, considering many exposition strategies, depict it in a story tree augmented with connections.

- Finish your design down to paragraph topic sentences before writing *any* of the draft. That is, *don't hack*.
  [Wait! This statement may seem extreme to you, but read on before concluding it's a bunch of hooey.]

- Draft one paragraph at a time.

- Edit in the small, not in the large. That is, produce a thesis in one draft — and smile.

Missing from this way of work is the bulk of the research, where you strive mightily so that you have something to write about.

How will you benefit from following this way of work? Well, you will avoid false progress, and will more efficiently produce a better thesis than if you had hacked.

Audience and products. Remember, you always write using words, but the audience varies by product. The table below summarizes the product-audience pairing. Yes, of course, your sketched or polished contributions and design will appear in the thesis, but, when you craft them before writing the draft, you are the audience. With a statement of contributions you say to yourself "This is what I intend to contribute". With a design you say to yourself "This is my plan of how I intend to inform them (my committee) of my contributions and my evidence for those contributions". With a draft you say to them "Look, read this. I have created something new and valuable, that somebody cares about, and I can prove it".

		Audience
	Contributions	*You*
Products	Design	*You*
	Draft	*Them*

Work backwards from contributions. Working backwards from contributions means three things. First, you articulate your contributions early for yourself. Second, you keep this contribution destination in mind, so you can more readily make design decisions, deciding what to include or exclude (bound the thesis), and deciding how to present your material. Third, you do not work forwards from what you could write. Remember, it's not about you; it's about them.

When you design your thesis for yourself imagine that in the draft you will be telling a story to them. The point of the story is your set of contributions. The context for the story is the state of your field without your work. To tell the story you need to work on two pieces: the vocabulary, and a story tree.

Vocabulary.   Vocabulary is an important structural component of your thesis. Because you live 24/7 with your thesis, you can easily forget that they do not, and then be surprised when they don't use vocabulary as you do, or don't know what your vocabulary means. Therefore, you should craft precise definitions for them early in design, and feature them prominently in your thesis.[11] One kind of vocabulary is the content terms that come from the field, or that you invent. (I defined "draft" early in this paper.) In my database work I teach people two phrases to help them precisely define vocabulary: "What do we mean by one of these?", and "Anchor your understanding with instances". These phrases might help you too.

A second kind of vocabulary is glue, the words, sentences, and paragraphs that make your writing flow for the reader. Poor glue words cause trouble. For example, in works I read, a writer often will use glue words such as method, methodology, process, and so on. Some writers intend them to be synonyms, and use the different words for variety. Others mean them to be different things, but don't explicitly say so. My advice: a) don't use synonyms for variety, pick one and stick with it, because the repetition tells the reader that you are referring to the same notion again; and b) explicitly define terms, because otherwise you risk the audience misunderstanding you, and thinking less of your contributions. Also, by paying attention to vocabulary you can minimize "drift". In something as big as a thesis, written over several months (or years☹), style and vocabulary drift will naturally occur. You need to take care to avoid it. Finally, you should consider the possible advantages accruing to most of those who generally avoid the overuse of somewhat qualifying words. That is, don't use weasel words (which I did in the last sentence.).

It behooves you to work hard at vocabulary, because otherwise you risk readers thinking less of your contributions.

---

[11] Voltaire supposedly said "Before you talk to me, define your terms".   "Supposedly" means I don't have the reference. Sniff, sniff -- sorry, your committee will frown on "supposedly said" quotes.

Figures. To help the reader perceive the meat of your thesis you should liberally use figures (including graphs and tables.[12]) Everybody uses figures for raw experimental data and statistical results (and I'll ignore such figures here), but figures can do much more, showing the audience what is coming and how pieces relate. For example, before you begin to design your literature review build a table. In the Appendix Figures A1 shows literature review summary by Libby Shoop. Each row is about a relevant research project. Each column is about a relevant property. Each cell in the matrix describes something about a project – property pair. You, of course, will have to decide what is relevant for your work – and doing so is good for you. The terms you use should be part of your content vocabulary. Also, pay attention to the ordering of columns and rows. Pick orderings that make it easy for you to tell your story.

Other large-scale figures might help both you and your readers. Figure A2 shows the field before and after Leone Barnett's thesis. Figure A3 shows a system diagram depicting what Jake Chen built and how it fits in bioinformatics. Each became an organizing principle for at least a thesis chapter.

Do not just toss in a figure and then ignore it. A well done figure provides an organizing principle for your writing. It gives the reader a 2-D picture to go with your words, which are inherently linear.

When you use a figure well good things happen:
- You will more readily see what you need to write about.
- You will see more to write about than you would have otherwise.
- You will make your points with more consistent vocabulary.
- The figure will suggest an organizing principle for the story contained in the figure.
- The audience will think more highly of your work than they would have otherwise.

---

[12] Check out R. Harris, "Information Graphics: A Comprehensive Illustrated Reference", Management Graphics, 1996.   I find it a valuable resource.

Spend time polishing your figures; it is well worth the effort. Figures A4 and A5 show an intermediate and a polished version of a figure from a research paper. The latter more clearly distinguishes between active (oval) and passive things (box). It also has less ink[13] and is easier to follow because related things appearing closer to each other allows shorter, no-bend lines. Figure polishing is helped by writing. So when you put a figure in your thesis, be sure to write about it, that is, make the figure part of your story. For each one, finish this statement: "Here is what you should get out of this figure …". Don't depend on the reader working hard to determine what is important. It is your job to do that work. As you try to tell a figure's story, you probably will find ways to improve both the figure and the story. That is what happened with Figures A4 and A5. When Bailey and I talked about A4, we found, to his surprise, that it did not support the paper's story, and that the story was incomplete. We fussed with A4 saying "What does this symbol connote", "What do we want to say here", and "What's missing". We fussed with discipline, and Figure A5 was the much better result.

Read other theses looking for figures that you can mimic. Pay attention to how well integrated they are with the text.

By the way, fields vary in where they tell a figure's story. Some, by convention, explain a lot in captions, while others don't, and, instead, tell the story in the body of text. Either way works fine.

---

[13] E. Tufte, "The Visual Display of Quantitative Information", Graphics Press, 1983. Another valuable resource.

Story tree. The heart of your design is a story tree, a detailed outline depicting the pieces and subpieces of your thesis, the order in which they appear, and the smaller signposts, the glue that guides the reader along the way.

Why must you concern yourself with designing a story tree? Well, your thesis undoubtedly contains an intricately connected network of topics, but readers read linearly. Somehow you must cover the network, and you necessarily will refer to a topic more than once (but, of course, not too often). With the design you specify the order in which the readers will see the topics, how much content you intend to present each time a topic is touched, and what you will say again, but with a different emphasis or elaboration.

Why use the term "story tree"? "Outline" is too vague, and could lead you to merely sketch out a few points and begin hacking away at a draft. "Table of contents" connotes too much the order in which the finished product appears. It's *post hoc*. However, "story tree" puts the emphasis on designing for readers. After all you tell a "story" to others, and therefore you think about what they already know and believe, and how you are going to get them to your destination for them — understanding and accepting your contributions. A "tree" connotes that you organize your story.

How detailed a design? This is crucial:

- **Design down to paragraph topic sentences.**

You should feel free to augment your PTSs or any part of the design with sketches noting the main points you intend to make, or key vocabulary terms. After all, you are the design's audience.

What does a story tree look like? It can take many forms, on index cards as McPhee works, or a word processors outline mode, or whatever works for you. However, it's the content, not the form that matters.

26

On the left below is a partially delineated skeleton of a story tree with a maximum depth of 4, and four placeholders for PTSs. In a real one you would put words with each oval, and where each "PTS" appears. While you surely would use at least a coarse outline anyway, remember that this story tree, when complete, is detailed down to PTSs. (You might prefer to rotate the tree 90 degrees so that it looks more like a table of contents.)

A story tree has another component: connections. On the right below is the same skeleton augmented with a few possible places for connections: I for introduction, T for transition, S for summary, and C for crossing. A crossing shows where, elsewhere in the tree, it crosses to.) You won't write connections everywhere, but you should explicitly decide, and not put connections where it happens to feel appropriate as you write.
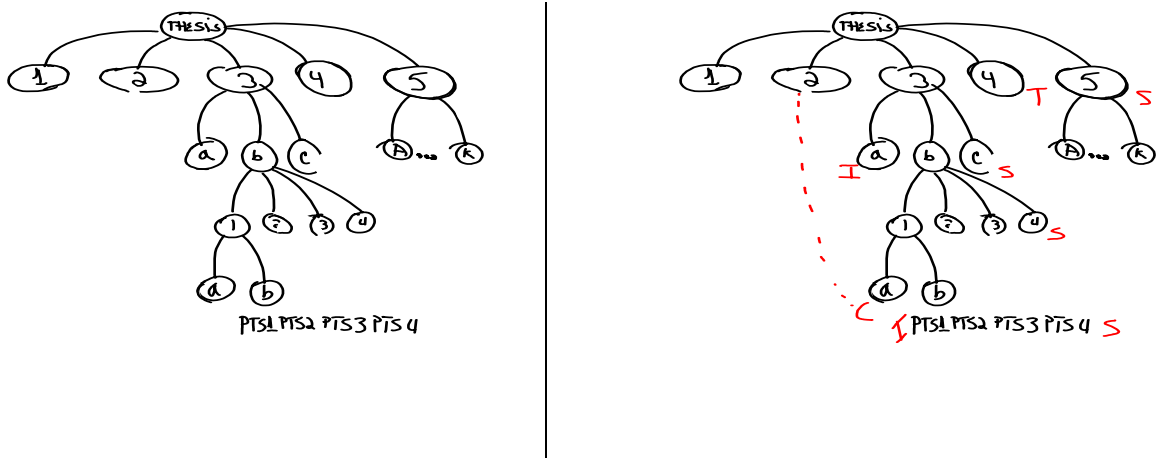
To help you judge the goodness of a story tree, use two notions that come from software: coupling and cohesion. Coupling measures the degree of interdependence between two pieces. Higher coupling means that if you change one piece then you more likely will have to change the second piece. Cohesion is the degree of relatedness of the notions within a piece. A piece with high cohesion is about one thing. (Look up "unity" in a writing handbook.)

Connection Discipline. Building a story tree helps you discipline yourself. You can't just say anything anywhere. The trick is to a) make each piece inward looking, that is, independent (or nearly so) of the other pieces, and b) limit connections between pieces,

especially Crossings. When thinking about connections, look at how two story tree pieces are related. They can have a sibling, a parent/child, or a cousin relationship. (The sections of a chapter are siblings; a section and its chapter have a child/parent relationship; a paragraph in section 3.b.1.b has a cousin relationship with anything in chapter 2.) You should expect to write I, T, or S connections showing how siblings or a parent-child pair connect. However, limit cousin connections — these are Crossings. You will need crossings, but don't put them among siblings; put them before or after.

How should you go about designing a story tree? Well, any real design process is messy, creative, and iterative, not top down, bottom up, or some other mechanical process. Don't worry, there are many paths to success, and, (I am serious about this) since you are an unusually smart person, you will find a way.

While there is no one best way to design, here are some things to try. First, experiment with organizing principles. For example, Libby Shoop took the readers on a tour of a system figure (not shown), going from the upper left, down, right, and up. Instead she could have reversed that order, or started in the middle and moved out from there. Let me say this stronger: build figures early because they will suggest organizing principles to you. Second, look at siblings and decide if they are a set or a sequence of topics. For a sequence, usually you will choose to first to last, or the reverse. For a set of siblings, any order will due, but you may order them most to least important, or the reverse. You might also choose an order that allows you to define terms first. Third, think about the balance of the tree. If one branch is large relative to others there may be no problem. However, you may decide to prune the large branch, and put its detail in an appendix. Fourth, to help you decide about pruning, keep the reader in mind and refer to your contributions. Fifth, after you design, put the story down, let it cool off, and come back to it with fresh eyes.

Finish your design down to PTSs before drafting. Here is a crucial rule:

- **Don't write the draft until your design is done.**

Said a second way:

- **Design the hard parts before you write the easy parts.**

Said a third way:

- *Don't hack!*

"Hack" is an interesting word. A hack is a) a person hired to do routine, often dull writing, or b) a bad thing (a hack - a failure), or c) a good thing but only in the short term (e.g., a hack job - duct taping a plastic sheet over a car's broken window). When you hack you cobble together a quick, temporary fix to a small problem. This rule adheres to a long held principle for software programmers: "design, design, design, then code once". In other words: *don't hack*.[14]

False Progress. Following this finish-the-design rule allows you to avoid the waste that comes from *false progress*. Many students, even those who know better, violate the rule in order to show "progress". However, real progress does not occur if you write easy paragraphs, and then either discard them when they do not fit, or keep them when they do not truly fit in the next draft. What does occur is either wasted effort, or diminished thesis quality. Since either outcome hurts, you should be patient, and not opt for false progress. If a paragraph is easy to write now, it will be easy to write later. You benefit if you were patient, and the PTSs, for the paragraphs that you didn't write, get moved, changed, or discarded.

False progress occurs a lot in software development. I know people who were told by their bosses to start writing code even though the specifications were barely begun. The bosses

[14] See John Carlis, "Wizards See and Dance, While Hackers Grope and Grunt", National Teaching and Learning Forum, April, 2000, (www.ntlf.com). Also see Patrick Starr and John Carlis, "Introducing the Design Process to Beginners: The Spiral Model", ASEE 2000, Learning and Teaching in the 21st Century, September, 2000. (www.cs.umn.edu/~carlis) Both articles address teaching design.

wanted to get going, but caused problems. Not surprisingly that software was tossed. You may have a similar problem with your advisor. (More on that in a bit.)

False progress occurs when you say "I don't know what I have to say so I'll just start writing", and fall into the wrong audience error, or fail to focus on contributions. Be careful to not misunderstand me. I write to learn, and recommend that you do to, but when I am noodling I am writing for *me*, not them, so I sketch.

You will be willing to be patient (i.e., not draft) if you can "see". See what? See the flesh that will be on the bones of your design, that is, the paragraph that delineates the idea of each PTS. What enables you to see? Mastery of your craft does.

Draft one paragraph at a time. If you have a quality design, with high cohesion and low coupling, then you can, and should, write one paragraph at a time. Doing so has several benefits. It allows you to focus on just one thing, and therefore makes you efficient. It raises your gumption level, because, while writing a *thesis* is a daunting task, writing a *paragraph* is do-able by a regular person. It keeps you from babbling, because all you do is develop the PTS. It serves as an alarm mechanism, because, if you are thinking about parts of the thesis other than the paragraph you are currently writing (plus, perhaps, its nearby paragraphs), then either you need a rest, or you have a flawed design that you need to fix.

In what order should you write the paragraphs? Good news: many different orders will get the job done —after all you created a quality design. Of course you should write term-defining paragraphs before any that use the terms, because otherwise you are not grounded and will waffle. Write a section all at once, rather than skipping around. Write paragraphs with relatively uncertain terms or PTSs early, so, if needed, you can revise the design. (Don't let this happen often.) Consider writing the meat of a section before writing the transition paragraphs. Write introductory paragraphs late.

Edit. When you produce the draft you will edit. Here is really terrific news: because you designed before drafting, most of your editing will be in the small, that is, within a paragraph, and not in the large, that is, you will not need to make structural changes. So your editing is merely polishing, not rebuilding.

Some loose ends.

Help!   What if part way through drafting, you discover that the design, and therefore the draft, is seriously flawed? Well, with house construction, you don't hack at the house. Instead, you consult the blueprint, fix the design (perhaps sketching alternative solutions), and then resume construction. You regret your bad planning and vow to do better next time. Do likewise when drafting a thesis.

What if you are stuck, and don't know how to tell a portion of your story? Here are several things you could try. Draw a figure or several. Study your vocabulary; perhaps it is vague or incomplete. Talk it out with your advisor or a friend. As I talk with one of my students when she or he is struggling, I often will hear a gem. When that happens, I say "don't talk anymore until you write that down". From such gems a better story emerges.

What if you just hacked and wrote an unacceptable product? Instead of hacking again, back up and design, finding a story.

What if your field, or advisor, has a standard template of 6 chapters, so you decided to begin with just that template.   Well, that's way too coarse an outline, and is like saying to a builder no more than "Build me a three-bedroom, two-bath house". Disaster certainly will follow. Sure, use that template if it fits, but do not think it will allow you to avoid the hard work. Be prepared to fight against that template if it doesn't fit.

What if, when you write the draft, you find that your PTSs are not really PTSs? You might find yourself writing several paragraphs for a nominal PTS. If so you may just have been lazy, and did not finish your design. You might have PTSs not leading to paragraphs at all.

If so you may need some remedial training to learn basic paragraph craft. Remember, make paragraph your unit of thought.

What if your advisor approves your design but then wants major changes to the draft anyway? (To avoid getting into this pickle, be sure that you and your advisor share expectations.) Sorry, I don't have a magic pill for this problem. The good news is that a quality design will limit the effort to make those changes.

A Caveat. When can you begin to flesh out the design? In other words, do you have to wait until the entire topic tree is deigned before writing at all? Well, if a portion of the topic tree is unfuzzy and relatively isolated from other parts, then you can draft it, while continuing to develop other parts of the tree, or to accomplish the research for other parts. For example, you might decide to draft your literature review early. (You may have a literature review from a proposal, and deem its organization adequate for the thesis.) However, be careful. The more you draft early based on an incomplete design, the higher the risk of false progress when you discover its flaws.

Size Limit. A thesis has no page limit, so do not worry about it. However, if you are writing for a journal or conference you will face limits. In the latter case use your topic tree to help you prioritize topics for pruning.

The Abstract. The thesis abstract is for people who may not read the rest of the thesis. It should be specific, stating the problem, your approach, and your contributions. A good abstract can motivate potential readers to dive into the thesis, because the impact of your contributions excites them. Don't make the mistake of making your abstract into a table of contents. For example, don't say "Next I present several tables showing a summary my experimental data". Such statements are about form not content.

Advisor. Where does your advisor fit in this way of work? In the figure below, your advisor (hatted) is privy to your sketches, and can (and should) give you feedback much earlier than the draft. Ideally your advisor approves the story tree, and then accepts the

draft you write from it. For that to happen the advisor also must be able to imagine the flesh that you will put on the bones. [An aside to advisors: I work this way, and get a better draft because I was in on the design. Try it, you'll like it.]
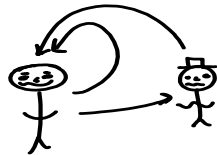
Celebrate, not defend. When you are done drafting and polishing, you must "defend" the thesis. However, you don't have to get defensive. You can turn your defense into a celebration, welcoming you to the community of scholars. How? Have good contributions, a well-written thesis, and a thesis talk that focuses on contributions. Your figures will provide an organizing principle for your thesis talk, just as they did for the written thesis.

Conclusion.

To help you believe in my design way of work, think about this: I can readily juggle three rocks, and my skilled friend John Riedl can keep six up in the air at the same time. Nobody can juggle a hundred rocks – vertically. However, horizontally (on a table) juggling a hundred is not a problem, because a rock stays where I put it and I can focus on one or a few at a time. Drafting without a design is like vertical juggling, and with one is like horizontal juggling -- much easier.

Finally, "joy" is important. Completing a Ph.D. is a long haul, and you need joy to sustain you. If you work on a topic that you have an abiding interest in, and follow the "Design a Thesis" way of work, then you can achieve success with joy.

Appendix

Table 2.1: Charactersitics and Use of Similarity Algorithms

Reference	Computation Method	Time Complexity	Space Complexity	Alignment Type	Typical Use	Scoring Measure (similarity or distance)	Allow For in Scoring Measure: indels/gaps	gap penalty for unmatched ends	other	Gap Weighting Function (e.g.: single, multiple, affine linear, affine concave)	Ability to Assess Statistical Significance
Needleman & Wunsch, 1970	dynamic programming	O(MN)	O(MN)	global	alignment of two presumably homologous sequences	similarity	yes	yes		single indels each penalized	Difficult
				local	alignment of common sub-sequences of two sequences which may have the same functional site	similarity	yes	no		single indels each penalized	Difficult
Sankoff, 1972	dynamic programming	O(MN)	O(MN)	local	alignment of common sub-sequences	similarity	yes	no		given number of indels treated as one and given different penalty	Difficult
Sellers, 1974	dynamic programming	O(MN)	O(MN)	global	alignment of two presumably homologous sequences	distance	yes	yes		single indels each penalized	Difficult
Waterman, Smith, & Beyer, 1976	dynamic programming	O(MN²)	O(MN)	global	alignment of two presumably homologous sequences	distance	yes	yes		multiple - allow indels of length greater than 1 ('gaps')	Difficult
Smith & Waterman, 1981b	dynamic programming	O(MN²)	O(MN)	local	alignment of common sub-sequences	similarity	yes	no		affine linear function for gaps	Difficult
Gotoh, 1982	dynamic programming	O(MN)	O(MN)	global or local	alignment of two presumably homologous sequences OR alignment of common subsequences	similarity	yes	set parameters to determine		affine linear function for gaps	Difficult
Taylor, 1984	dynamic programming	O(MN)	O(MN)	global or local	alignment of two presumably homologous sequences OR alignment of common subsequences	distance	yes	set parameters to determine		allowed lower penalty for large gaps	Difficult
Waterman, 1984a	dynamic programming	O(MN) (est.)	O(MN)	global or local	alignment of two presumably homologous sequences OR alignment of common subsequences	similarity	yes	set parameters to determine		affine concave function for gaps	Difficult

Figure A-1 Part of Libby Shoop's literature review summary table.

Figure 1.1. Big Picture

Extending DBMS to Support Semantic Information in GIS
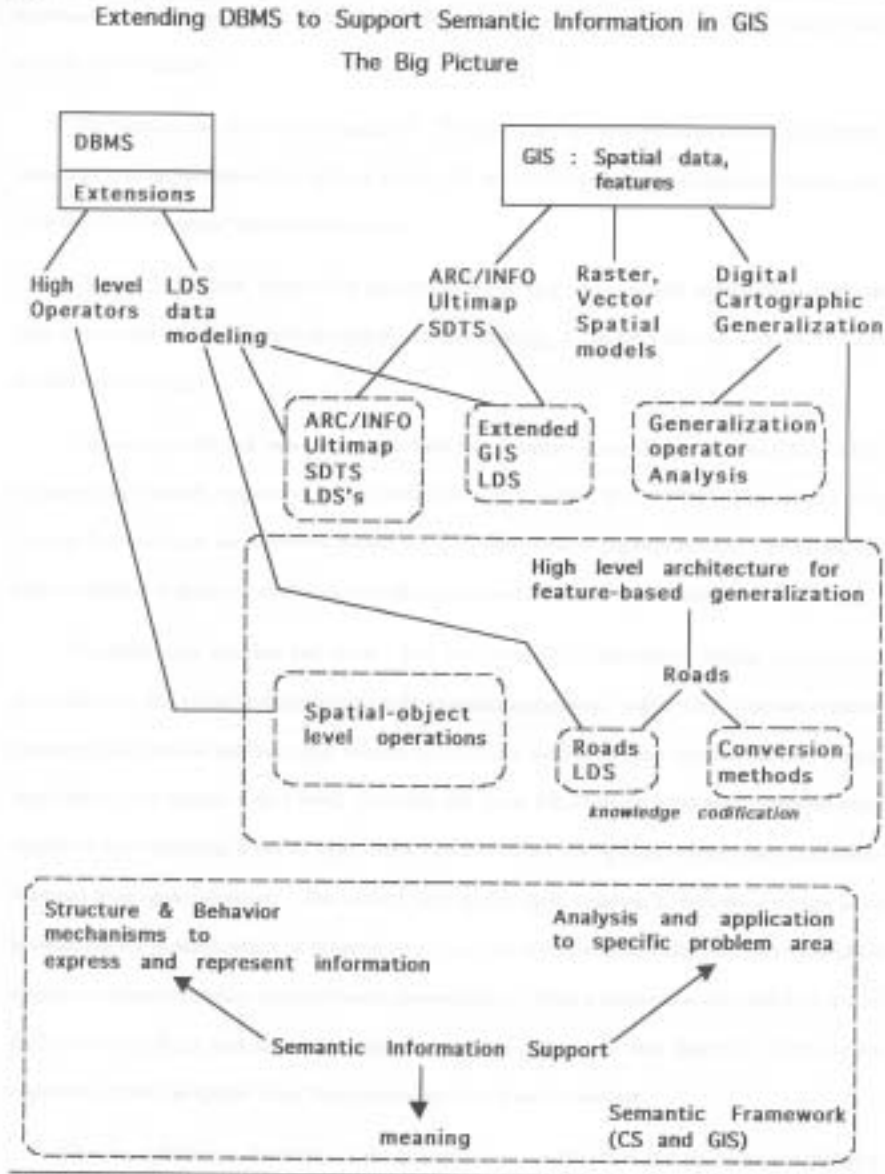
The Big Picture

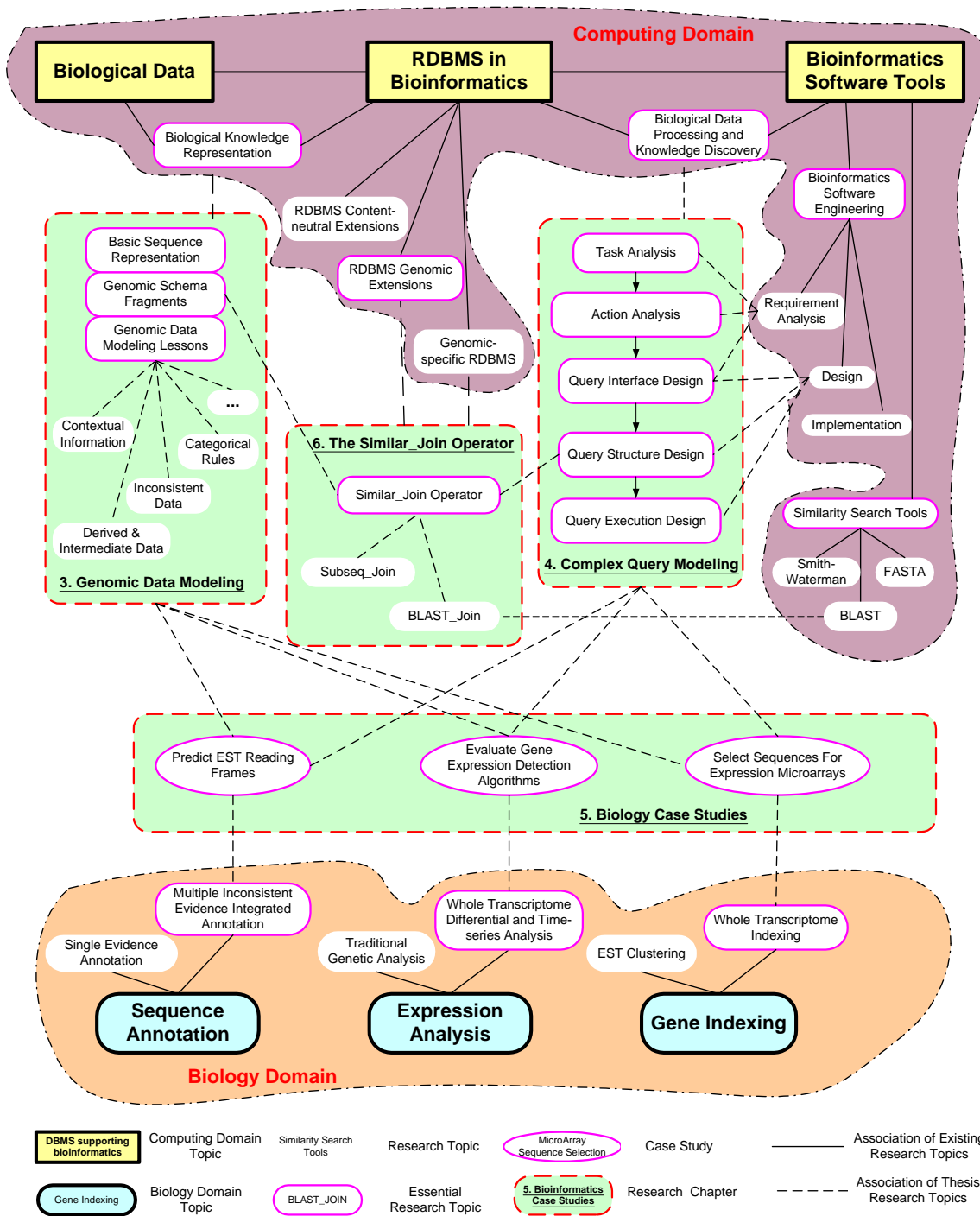Figure A-2 Leone Barnett's before and after figure.

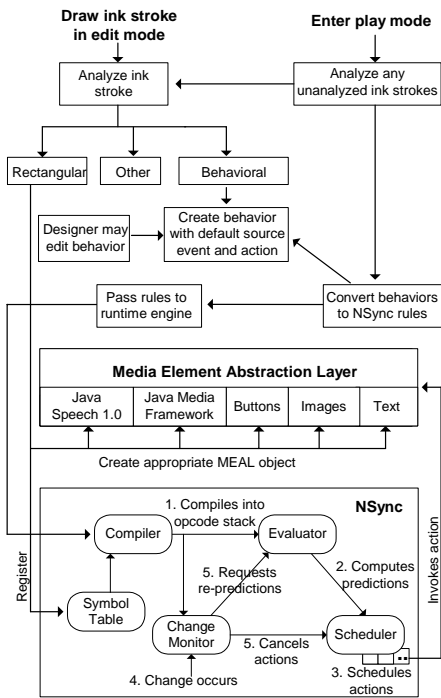Figure A-3 Jake Chen's Thesis Overview Diagram

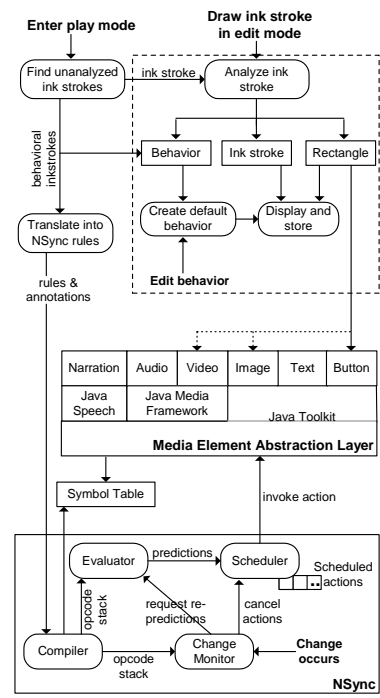Figure A-4   "Intermediate"

Figure A-5 "Polished"

Figures A-4 and A-5 are from Brian Bailey, Joseph Konstan, and John Carlis, "DEMAIS: Designing Multimedia Applications with Interactive Storyboards", Proceedings ACM Multimedia 2001.