Literature Reviews (First circulating draft)

Cem Kaner

Florida Tech

January, 2004

These notes are partially based on research that was supported by NSF Grant EIA-0113539 ITR/SY+PE: "Improving the Education of Software Testers." Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

What is a literature review?

- Organized examination and presentation of the writings in a field of study or practice.
- In the computing fields, we might study:
  - scientific (theoretical or research) publications (including experiments, case studies, models, etc.)
  - experience reports
  - educational publications (such as books that are intended to teach the reader rather than report new findings)
  - tutorials
  - marketing materials
  - product documentation
  - news reports of events of interest

Examples

- The introductory books' descriptions are very general, and the advice they offer is very generic, more applicable to some cases than others.
- Before considering general advice for doing a review, I want to present several very different examples of reviews that I've done (that served their purpose well).

Example 1: Fornix-fimbria lesions and spatial learning in rats

- CONTEXT: A recent literature review stated that a certain model of hippocampal/fornix function predicted no difference between lesioned and normal rats in a certain type of experiment, and summarized results of that type of experiment as showing no difference. I ran a variation of that experiment and found enormous behavioral/learning differences.
- QUESTION: WHY THIS INCONSISTENCY?
- FOCUS: (a) Methodological, such as intensity of shock, shape and other details of the maze, type of surgery, extent and precise location of brain damage, relationship-building activities between the experimenter and the rat, types of measures taken and reported. (This was exhaustively done.)
- FOCUS: (b) Related research, such as other lesions or other studies of spatial learning. (Sampling only, search for interesting ideas.)

Example 2: Psychophysical measurement models

- CONTEXT / GOAL: Understand the theory and mathematics of a set of models relating subjective intensity (such as perceived brightness or loudness) to physical intensity.
- FOCUS: the models (and the key papers describing the models).
- SCOPE: Data merely instantiated the models, methods merely instantiated the data. Gather enough of each to provide illustration, clarification

In Example 1, methodology was everything, theory was almost nothing. In Example 2, methodology is (almost) nothing, theory is almost everything.

Example 3: Reckless Driving

- Context: In California, to prove "reckless driving", you have to prove that the driver's state of mind was "reckless" -- essentially that she acted with conscious disregard for the safety of others. This requirement makes reckless driving the hardest misdemeanor to prosecute, in many ways more similar to a prosecution for murder than for other driving-related misdemeanors.
- In other states, it is sufficient to prove that the driver did 3 illegal maneuvers within a short period of time, potentially endangering others. Most misdemeanor prosecutions are handled by junior prosecutors and many of them try the 3-violations approach. The police collect evidence to support this style of prosecution. The problem is that this is not necessarily proof of recklessness under California law.

Example: Reckless Driving (2)

- I approached this a different way, reviewing court cases to determine what types of evidence judges considered suggestive of recklessness. I put these into a big chart for the cops (who now gathered additional information at the scene of the crime, providing a much better basis for prosecution), and wrote a detailed memo (listing the cases I relied on for each item in the chart, with additional comments) for fellow prosecutors to use in court.

Example: Reckless Driving (3)

- GOALS OF THE REVIEW: (a) Identify a large set of facts that courts are likely to consider relevant to determining the state of mind of the allegedly-reckless driver. (b) Provide references to the relevant court decisions, and point out any complexities involved in using a given type of fact.
- SCOPE OF THE REVIEW: about 1000 court cases. Most of these were not useful for the final review, but this is what I sifted through to get the 250 or so cases that I cited as references.
- EXTENT OF THE SEARCH: Exhaustive search for ***types of*** relevant facts. Non-exhaustive search for exemplars (persuasive or informative examples) of the use of the facts.
- *How much detail should I have captured for each case?*
  - Relevant details: (a) Reckless driving charge under CA statute; (b) a given type of fact was discussed or mentioned and accepted or ruled on as relevant; (c) the discussion or ruling; (d) the level of the court (Supreme Court, appellate court, etc.)
  - There were many additional details per case, but I decided they were irrelevant for this review.

Example 4: Freud's Theory of Human Nature

- CONTEXT: 3-essay review, for a 6-credit 4th year psychology course on theories of human nature.
- GOALS OF THE REVIEW: Understand Freud on his own terms, through a thorough analysis of his writing.
- SCOPE OF THE REVIEW: All of Freud's writings were candidates for review. Historians' presentations of background information on Freud were of interest but were sampled, rather than (as in Freud's writings) almost exhaustively searched. Other psychologists' reactions to Freud's work were typically out of scope.

In the first three examples, we consulted many different authors. In this example, we focus on one author.

Example 5: Legislative analysis, Uniform Computer Information Transactions Act

- CONTEXT: Political debate over a statute that had been in drafting process for 12 years.
- GOAL: Create an archival, authoritative summary of arguments in opposition, explaining the implication of this bill for the practice of software engineering. Explain legal issues to engineers, engineering issues to lawyers.
- SCOPE: Exhaustive review was impossible (huge literature), but issue by issue, presented cases or examples that clearly illustrated each point (plus important qualifying or contradicting cases or arguments). Selection of issues to cover was the hardest part of the review.

In this example, the selection of issues was not driven by the literature but by the author's understanding of what would educate the target readers. The literature provided detail, but not guidance as to choice of topic.

Example: Modeling time perception

- CONTEXT: I ran (and replicated) an unusually large study using a common method for asking people how long (in time) a given stimulus appeared to be. I replicated the basic results, but with the large set of data, I could analyze for finer-grain effects. These were inconsistent with the most widely accepted models of how this method worked.
- Over 2.5 years, I revisited the relevant literature at least 6 times with new questions on timing theory, math and physiological models, this methodology and related methods, response biases, and so on.
- The key point to note is that the same papers carried entirely new meanings or information for me when I reread them later from a new perspective. A "thorough" abstract of the paper in the first year would not have provided the information I needed later.

Example: Search & Seizure law

- CONTEXT: As a prosecutor in Drug Court, I often faced hearings on the legality of a search (that found drugs). In each case, I had to provide a literature review (a "brief") explaining the application of the relevant court cases to the particular search (of the person or property) and seizure (of the drugs, guns, etc.) at hand.
- MATERIALS: I kept a set of binders, with all the key cases sorted chronologically. The cases had Post-It abstracts (lists of key words on a Post-It note on the first page of the case). For each hearing, I pulled out the relevant cases, put them in a hearing-specific binder, summarized the relevant literature, and brought the binder and the summary to court.

In this case, the document collection is everything, the summaries are of minor lasting importance.

Example in progress: Domain testing

- CONTEXT: The testing literature talks about this technique as if it were uncontroversially defined, yet there are inconsistencies in description that lead to differences in practice. The typical descriptions of the technique are either shallow (and applied only to single variables) or nearly incomprehensible (when applied to multiple variables).
- GOAL: Identify the different visions of domain testing, the classic papers that describe each or provide detail or data or application guidance for them, figure out what the actual commonalities and divergences are, and create comprehensible applications of individual strands (rather than of the internally inconsistent lot) to multiple variables.
- GOAL: Provide teaching support materials that take a mature look at domain testing (explaining the divergences in a clear way).

Example in progress: Scenario testing

We're figuring out:

- what are the key attributes of this technique
- how to design excellent individual scenario tests
- how to research and design a scenario testing series or strategy

Challenges:

- Multiple meanings of "scenario" confuse readers
- Much of the insight into the value of scenarios was gained in other disciplines, often with other vocabulary.

Example in progress: High volume test automation

- CONTEXT: This is a generalization from about 10 different test techniques used in industry (and in many cases, not much described in the academic literature).
- CHALLENGES:
  - What ARE the sources? Fragmented literature, non-overlapping discussions of related methods, fragmented vocabulary.
  - Connecting to the literature and reading preferences of a scientific (CS) community that seems largely clueless and happy to stay that way.
  - Need examples of applications
  - Need published details so that we don't have to rely on information that is under nondisclosure in order to explain the method and its use.

The review process

- Cooper (p. 5) lists 5 key "stages" of a literature review. Rather than pretend these are done in a specific order, let's call them key activities, which might be done in parallel or in any order:

" (a) problem formulation;

(b) data collection or the literature search;

(c) data evaluation, in this case assessing the quality of studies;

(d) analysis and interpretation; and

(e) presentation of results."

The review process

- Each of these activities is important.
- In this talk, I'll spend more time on the goals of the review, problem formulation, doing the literature search, and analysis & interpretation.
- My notes are informed by Cooper's descriptions of these activities, but differ significantly from Cooper's descriptions. Cooper is an interesting alternate source.

Problem formulation

- Includes everything from deciding what topic you want to write on to deciding what information you want to collect--deciding what is *relevant* to the problem or question you are writing about.
- I assume in this talk that you have a basic objective in mind already, that you are doing a literature review because you need the information for a purpose, rather than because you have to choose an essay topic for a course and don't know where to start.
- So, suppose that you have a reasonable starting point, such as:
  - How could we exploit ambiguities in natural language specifications in order to design more powerful tests?

Problem formulation

- Let's break this down:
  - How could we exploit ambiguities
  - in natural language specifications in order to
  - design
  - more powerful tests?
- So what do we know (what can we learn about what is known) about **ambiguity**?
- What can we learn about **natural language specifications** ?
- **What can we learn about test design**?
- What can we learn about the **power of tests**?

Problem formulation

- At some point, we'll probably prioritize. For example,
  - maybe we only need a good understanding (rather than a thorough review) of the concept of powerful tests,
  - maybe the research in natural language specification is impossibly large and so we have to restrict our attention to those aspects of what is known about natural language specifications that might help us learn about ambiguity or testing
- At ***some*** point we'll do this because we need to narrow the topic to something we can actually do. But it might take us a while to understand the literature well enough to be able to intelligently prioritize and focus. This might come after a lot of progress in the other activities has been achieved.

Problem formulation

- To develop a feel for the problem, I want to learn what are the X's (see list below) of interest. X:
  - Variables that might be manipulated (or occur naturally) in research of our topic
  - Questions or issues that have been frequently discussed in connection with our topic
  - Disagreements among scholars or practitioners
  - Sources (key people, books, journals, articles, web sites, electronic indexes)
  - Keywords (terms that I might search under)
- Probably, if they are available, I'll start this by reading the "classic" introductory books, articles, or encyclopedia entries. "Classic" might not mean historically most important, it might mean the ones best used for teaching the subject.

Problem formulation

- Ultimately, I might have answers for questions like:
  - What search engines should I consult (and how many)?
  - Who publishes on this topic?
  - What is the vocabulary in this area? Is it consistent or do different groups use different terms for the same things and the same terms for different things?
  - How fragmented (lack of cross-reference to related work by another group) is the literature?
  - What bias should I expect from a publication (based on author, school, journal, etc.)?
  - What research paradigms (key ***types*** of experiments, perhaps in different fields, and the theoretical context around them) are used/useful for this topic?
  - What theoretical contexts does this topic arise in? (What theoretical implications have people seen for this topic?)
  - What are the controversies and confusions in this area?
  - What issues (content issues, such as, what types of ambiguity are there) seem important for an understanding of the topic?

Problem formulation

- At this point, these questions guide me in developing a research strategy. They aren't the answers. They are the questions I want to explore in order to figure out how to look for the answers to my underlying question:
  - How could we exploit ambiguities in natural language specifications in order to design more powerful tests?
- There are a lot of questions here.
- I do ***not*** assume that a reading for one purpose (such as, learning who is publishing in the field) will provide me with the information that I need for another purpose (such as, learning the vocabulary). I ***don't*** try to force my notetaking to cover all questions at once--it is too hard, too time consuming, too easy to lose focus, too frustrating (for some people, including me) and too likely to result in me giving up.

Problem formulation

- This is a critical strategic issue, and one where I differ fundamentally from authors like Garrard, Cooper, and I think Hart and Fink.
- I expect to read the sources several times and to use my notes as an index to the sources rather than as a replacement for (a summary that can be consulted instead of) the sources.
- My passes through the literature, to fill out charts or other organizers, are to help me build my understanding of the issue at hand, rather than of the fine details of the sources. It is natural for me to do many passes, one for each issue.
- (I'll still read the original source in enough depth to understand the context it provides for the information I'm pulling out of it, but the context is context, not essential content.)

Problem formulation

- RECAP:
  - I typically start with an opening library of secondary sources and/or educationally useful primary sources.
  - The opening library introduces me to the field and helps me develop a first-draft strategy:
    - WHO should I read?
    - WHAT QUESTIONS should I read to answer?
    - WHERE should I look for readings?
    - WHAT do I currently think is my ULTIMATE OBJECTIVE for this review?

Problem formulation

- Garrard (p. 116) says

  "In a review of the literature, the 3 most important decisions you will make are (1) specifying the purpose of the literature review, (2) selecting the source documents, and (3) choosing the column topics. Column topics in a review matrix are the issues or concepts used to abstract each journal article or other source document in a review of the literature."

  A critical 4th decision is the level of detail written into each cell (or into the article's abstract, if you use an abstract rather than a matrix). Garrard is both much more (a) generic and (b) detailed than I am. I want the shortest simplest summary in the table, pointing me back to the original source for elaboration. I write from the article, not from the table. It's like the "story" in XP, which is a one-sentence summary, fleshed out via conversation.

Problem Formulation

- Garrard (and others) advocates a thorough abstracting of each paper. Similarly, in law school, students are urged to write thorough abstracts (called "briefs" of each court case).
- Advantages:
  - powerful discipline for the novice reader, forces the novice into active reading, critical reading, and a systematic analysis. Educates the novice in the key concepts and methods and reasoning of the field.
  - a good summary jogs your memory and saves your need to reread things you've already read, understood, and summarized.
  - in theory (but I think not so much in practice), a thorough abstract saves you the need to reread later when exploring other issues.
- Disadvantages
  - Burdensome
  - Violates YAGNI, you often summarize for the wrong future issues and miss the issues that will be critical

Literature search

- In software testing, Florida Tech has an unusually broad collection of books. We are in the process of ordering several more, especially dissertations. If we don't have the book, we can get it via inter-library loan.
- Florida Tech has an extensive set of online resources for engineering research, such as the ACM portal, the IEEE journals, etc. The library offers a research seminar every term to familiarize students with the online resources. If you haven't taken it, you should, this term. It is free. I will not reiterate any of the contents of that course in this presentation.

Literature search

- One issue for search is the list of keywords to search under. Our literature is often fragmented.
  - For example, in the area of high-volume automated testing (HVAT), the literature discussing HVAT techniques for telephone systems might have zero overlap with literature on HVAT for medical devices.
  - As a broader and more distressing example, the academic literature and the practitioner literatures overlap minimally.
- You might find the same concept referred to under five different words, in five different areas of research or application. Rather than deciding what your favorite name for a concept is, and sticking with that one, your search should probe the field under each name in use.

Literature Search

- Garrard lays out 4 artefacts that are important to create:
  - The "paper trail"
  - The "documents section"
  - The "review matrix"
  - The "synthesis"
- The paper trail is a record of the research sources you've consulted so far. For example, suppose that you are searching under 8 terms, and you are searching the ACM Portal, the Engineering Village, CiteSeer, ERIC Digests, ProFusion, and Teoma. When I do this, I make a chart (next slide).
- This tracking is essential or you will miss some searches, believing that you searched source X for topic Y when you had not.

Literature search

	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6	Word 7	Word 8
ACM Portal								
Eng Village								
CiteSeer								
ERIC								
ProFusion								
Teoma								

Inside the matrix, the cells show the date you did the search and/or other key notes.

Literature search: Documents section

- The "documents section" is your virtual binder of all the materials you are citing. I use hardcopy binders or organizer files. Garrard suggests sorting by date. I prefer sorting by author name. In either case, I would stick post-its on each document to show its keywords.
  - Make YOUR OWN SET of copies
  - MARK THEM UP freely
  - Include photocopies of directly relevant sections of books (with their bibliographies)
  - Sort them by issue on an ad hoc basis: pull everything you have on a given issue, write about the issue, re-sort back to original sort order.
  - Build an index to your library (EndNote)

Literature search: Matrix

- The matrix is the core of your literature research. What should it contain? Avoid formulaic template-like answers.
- Column should include author(s)
- Probably column for date (might merge with authors column)
- Might have column for title (I don't, it's too long)
- The rest are the issues you currently think are interesting. You are collecting data into these columns, fill them in (for a given paper) when that paper has something to say about the column's issue, briefly note it. Else leave it blank.

Literature search: My iterative process

- Start with a starting library (intro texts, encyclopedia, literature reviews, detailed descriptions. These yield vocabulary of topics/concepts and key sources (people, papers, journals/publishers, etc).
- Keyword searches
  - goal: find influential papers, find literature clusters
  - risks: easy to misclassify a relevant or irrelevant source b/c of mass of material that shows up in this search
- Work forward/backword through literature:
  - Read X paper's sources
  - Read papers that cite X
  - Big risk of fragmented literature, miss parallel set of work
- Collaborator reports (ACM Portal or CiteSeer)
- Literature reviews
- Dissertations
- Websites of key contributors

Literature search: My iterative process

Meta-questions

- I am constantly asking these questions:
  - What am I trying to learn?
  - What are the dimensions of what I am trying to learn?
  - What other literatures / fields are tied in with what I am trying to learn?
  - Where is the relevant information?
  - How can I assess credibility of sources?
  - Who is my audience and why are they interested and what do they want to (what do I want them to) learn?
  - What foundations do I need to understand this literature?

Evaluate Sources

- For experiments, there are lots of ways to assess validity of the experiment and the quality of the report. See, e.g., Girden's and Leavitt's books.
- There are rules of thumb for articles by source, such as the heuristic that articles in peer-reviewed journals are better than articles in non-reviewed or commercial journals. These are biases, heuristics, subject to error. In testing, several of the peer-reviewed articles are lower-value than the commercial articles. (But overall, the heuristics are probably more often accurate than not.)

Evaluate Sources

- For online sources, there are lots of heuristics and suggestions for assessing the quality of an online paper. Here are some of Garrard's suggested links:
  - Grassian, E., Thinking critically about World Wide Web resources, www.library.ucla.edu/libraries/college/instruct/critical.htm
  - Purdue University, Evaluating World Wide Web information, thorplus.lib.purdue.edu/research/classes/gs175/3gs175/evaluation.html
  - Teaching students to think critically about Internet resources, weber.u.washington.edu/~lbr560/NETEVAL/index.html
  - Bibliography on evaluating internet resources, refserver.lib.vt.edu/libinst/critTHINK.HTM

Writing the review

- The matrix columns lay out a set of issues. In effect, they are the outline for your paper (maybe not in the order you'll write the paper). Each section of the paper corresponds to a column, an issue that you now understand thoroughly, and you know who wrote what about it. Write the issue in your terms, what you are trying to teach people about the issue (with references to the sources for their facts, but no need to push their words).

"Owning the Literature"

- Garrard, p. 6, a person who successfully reviews the literature on a topic knows the area so well that they "own it."
- You know you're beginning to own the literature when:
  - you know what study an author will cite to make a pooint, before they cite it
  - you recognize the facts and can identify the study before they cite it
  - you spot plagiarism (of words or ideas)
  - you notice absence of citations to key sources and can think intelligently about whether this is due to sloppiness, fragmentation of the literature, bad blood between two researchers, or something else
  - you highlight an area of open questions or needed research and then see recent dissertations or conference papers or email discussions of it
  - you can predict the perspective and methodological approach of a paper from the author's name or institution

References

- Cooper, H. (1998). *Synthesizing Research*, 3rd Ed., Thousand Oaks, CA: Sage Publications.
- Fink, A. (1998). *Conducting Research Literature Reviews*, Thousand Oaks, CA: Sage Publications.
- Garrard, J. (1999). *Health Sciences Literature Review Made Easy: The Matrix Method,* Gaithersburg, MD: Aspen Publishers.
- Girden, E.R. (2001). *Evaluating Research Articles,* 2nd Ed., Thousand Oaks, CA: Sage Publications.
- Hart, C. (1998). *Doing a Literature Review: Releasing the Social Science Research Imagination,* London: Sage Publications.
- Leavitt, F. (2001). *Evaluating Scientific Research: Separating Fact from Fiction,* Upper Saddle River, NJ: Prentice-Hall.
- Pan, M.L. (2003). *Preparing Literature Reviews*, Los Angeles, CA: Pyrczak Publishing.