Writing a Literature Review

ICS690 Seminar

Grad. Chair: Henri Casanova
henric@hawaii.edu
Research Papers

- Lit review: what and why?
- How to read papers?
- How to write a lit review?
Novelty of Research

- (Good) research must be novel
  - After reading a research article/thesis, what have we learned that we didn’t know before and that advances the field?
  - What we know: all that’s been published before

- If not novel, your work isn’t research
  - Possibly beautiful, clever, useful, and marketable engineering
  - “But I spend 2 years writing a super useful piece of code using my amazing coding skills...”
  - What we learned: you’re a great coder. doesn’t advance the field

- If you’re doing a Ph.D. thesis, novelty is paramount
  - You just can’t graduate if your work isn’t novel

- If you’re doing a M.S. thesis, novelty is less of a concern
  - Could be implementation/evaluation of research results

The good news: doing something novel is fun and rewarding!
Novelty of Research

- To find a research topic and ensure its novelty, read a lot of papers in your field
  - read, read, read, find gaps, choose one, fix it
- You have to start reading papers early on
  - Nothing worse than working on a project for 6 months and then realizing that you’ve re-invented the wheel without improving it
  - Not as bad: realizing after 6 month that there was some published work that makes your life simpler
- No matter what, you must convince the reviewers of your article that your work is better than previous work
- Papers that don’t have a comparison (qualitative or quantitative) to previous work are basically rejected out right from reputable conferences/journals
  - Papers are rejected all the time because they fail to reference some relevant previous work!
What’s a Lit Review?

- To do good research, you must know the literature inside out, which means that you could write a good lit review
  - and that way you won’t seem clueless :)
- **Definition**: A survey of a body of knowledge, which *critically* evaluates and contrasts relevant published articles while highlighting their contributions and findings
Why Write a Lit Review?

- To write the “related work” section of a research paper, or the “related work” chapter of a Ph.D./M.S. thesis
  - So that you can highlight your own contributions and contrast them to the reviewed literature
  - Note that itemized lists of original contributions are always a very good idea in papers and theses
    - “In this paper our contributions are: 1) ... 2) ...”
- To write the lit review needed for your Ph.D. portfolio
- To publish a lit review
  - e.g., in the ACM Journal of Computing Surveys
Research Papers

- Lit review: what and why?
- How to read papers?
- How to write a lit review?
Reading Papers

Once you have identified a research area of interest, you can start looking for relevant published work.

First, come up with a list of likely keywords:
- Being broad in the list of keywords is a good idea to not miss anything (I speak from personal experience).

Where to find papers:
- Free resources: Google, Google Scholar, Citeseer, Authors’ Web pages.
- Journals, conference proceedings:
  - ACM/IEEE digital libraries, to which universities have subscriptions.
  - Finding the relevant journals/conferences for your area is key:
    - so that you can look at TOCs of all past proceedings.
Reading Papers

- You must read a lot of paper
  - Requires discipline
- Don’t be afraid to be broad
  - Learning new things, being exposed to ideas will always be beneficial in the long term
  - Much good research comes from combining ideas from different areas
  - Useful to define your research area
- Keep an annotated bibliography:
  - Keep track of the bibliographical information
    - bibtex entries are good
  - Write a short informal summary of each paper, with keywords
  - This will constitute an invaluable resource
Reading Papers

- You must **follow references up and down**
- Following references down is easy
  - Just look at a “references” sections
- Following references up used to be very hard, but now is easy as well:
  - Google Scholar (let’s look at it)
  - Citeseer
- Building a graph of paper references is a good idea
  - Mental graph is ok, but written down is better
Research Papers

- Lit review: what and why?
- How to read papers?
- How to write a lit review?
The Scope

- Defining the scope is known to be difficult
  - Not too broad, not too narrow
- No silver bullet method, just rules of thumbs:
  - Can you state in a complete sentence exactly what your review is about?
  - Try to cap the number of articles you reference, and when you get over that cap, narrow your scope
- It’s common to see the scope change in the process of writing the review
- Nowadays, there are 100’s of articles on many topics, so a broad scope is very hard to do
  - Article selection is key (use critical thinking)
The Audience

- A sophisticated audience of people in your field, who are not experts in the particular area

- After reading your review, a previously uninformed researcher should be able to:
  - Engage in an intelligent conversation about the area with other researchers, including those in the area
  - Have a good idea of what’s known and of what challenges and big questions remain
  - Come up with some possible research paths in the area

- A good lit review is an invaluable resource
  - You should hope to find one in your research area
The Writing Approach

- Don’t dive too deeply into technical details
  - Readers can always go hunt references if they want!
- Instead, give the essence of existing knowledge
- Instead of quoting, describe all previous research with your own words
  - Allows you to synthesize the research
  - Allows you to harmonize terminology
  - Provides a single “voice” for the lit review

- Plagiarism:
  - You can quote sentences and include figures
  - They must all clearly reference published work
The Writing Approach

- It’s very easy to lose the reader in a lit review
  - After all, this is supposed to be all new to the reader, and comes from many sources

- Provide the reader with “umbrella” sentences at the beginning of sections/paragraphs
  - e.g., “In this section we review those works that have proposed cache-oblivious algorithms for linear algebra kernels.”

- Provide “signposts” throughout
  - e.g., “We have seen that the work in [10] advocates for the use of spectral decomposition. By contrast, the work in [12] ...”

- Provide brief “so what?” summaries at key points
  - e.g., “Based on the results in [8,12,18,42], reviewed in the previous section, the success of an approach based on compiler-drive optimization seems unlikely at best.”
The Writing Approach

- Make sure your bibliographical references are all correct, complete, and all there!
- Make sure that they are sorted nicely
  - Alphabetical by Author is what most reader will expect
The Structure

- Standard structure:
  - Introduction
  - Body
    - Multiple sections
  - Conclusion

- Let’s say a few words about each section
Lit Review Introduction

- Identify the issue
  - What it is?
  - Why we should care?
    - What are the applications
- Point out overall trends in what’s been published so far
  - Major conflicts
  - Major methodological differences
- State how the lit review is organized
- State why some literature is not included
The most difficult task is to organize the body of the lit review. One must categorize published works. The goal is to find common denominators to group some works into a single category, and to find discriminants to contrast the categories. Some authors try to come up with a hierarchical taxonomy. Not all powerful, and, taken to the extreme, an exercise in futility. Some authors present large feature/property tables. Let’s see some examples:

Hierarchical Taxonomy [1]

Fig. 3. Classification of top-\(k\) query processing techniques.
## Feature Table

### Table I. File system classification

<table>
<thead>
<tr>
<th>Feature</th>
<th>Status</th>
<th>Authentication</th>
<th>Authorization</th>
<th>Granularity</th>
<th>Autonomous Delegation</th>
<th>Revocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFS</td>
<td>P</td>
<td>AUTH_SYS, Kerberos</td>
<td>ACL (UNIX)</td>
<td>File system</td>
<td>No</td>
<td>ACL</td>
</tr>
<tr>
<td>NFSv4</td>
<td>P</td>
<td>Kerberos, LIPKEY, SPKM</td>
<td>ACL (NT)</td>
<td>File</td>
<td>No</td>
<td>ACL</td>
</tr>
<tr>
<td>AFS&amp; Coda</td>
<td>P</td>
<td>Kerberos</td>
<td>ACL (AFS)</td>
<td>Directory</td>
<td>No</td>
<td>ACL</td>
</tr>
<tr>
<td>CIFS</td>
<td>P</td>
<td>Plaintext password, Challenge-Response, Kerberos</td>
<td>ACL</td>
<td>Directory</td>
<td>No</td>
<td>ACL</td>
</tr>
<tr>
<td>xFS</td>
<td>E</td>
<td>AUTH_SYS, Kerberos</td>
<td>ACL (UNIX)</td>
<td>File system</td>
<td>No</td>
<td>ACL</td>
</tr>
<tr>
<td>Truffles</td>
<td>E</td>
<td>Public Key (X.509)</td>
<td>ACL (UNIX)</td>
<td>Volume</td>
<td>Limited</td>
<td>No</td>
</tr>
<tr>
<td>Bayou</td>
<td>E</td>
<td>Public Key</td>
<td>AC Certificate</td>
<td>Data Collection</td>
<td>Limited</td>
<td>Revocation certificate</td>
</tr>
<tr>
<td>WebFS</td>
<td>E</td>
<td>Public Key (X.509)</td>
<td>Hybrid</td>
<td>File</td>
<td>Limited</td>
<td>ACL, CRL, OLA², Certificate Expiration</td>
</tr>
<tr>
<td>CapaFS</td>
<td>E</td>
<td>No</td>
<td>Capability</td>
<td>File</td>
<td>Limited</td>
<td>CRL, Timeout</td>
</tr>
<tr>
<td>SFS</td>
<td>E</td>
<td>Public Key</td>
<td>ACL (UNIX)</td>
<td>File</td>
<td>No</td>
<td>ACL, CRL</td>
</tr>
<tr>
<td>GSFS</td>
<td>E</td>
<td>Public Key</td>
<td>ACL (SFS)</td>
<td>File</td>
<td>Limited</td>
<td>ACL, CRL</td>
</tr>
<tr>
<td>DisCFS</td>
<td>E</td>
<td>Public Key</td>
<td>Trust Mgmt. Credential</td>
<td>File</td>
<td>Yes</td>
<td>Credential Expiration</td>
</tr>
<tr>
<td>Technique Used</td>
<td>Section</td>
<td>References</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------</td>
<td>---------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixture of Models</td>
<td>Section 7.1.3</td>
<td>Byers and Raftery [1998], Spence et al. [2001], Tarassenko [1995]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression</td>
<td>Section 7.1.2</td>
<td>Chen et al. [2005], Torr and Murray [1993]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bayesian Networks</td>
<td>Section 4.2</td>
<td>Diehl and Hampshire [2002]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support Vector Machines</td>
<td>Section 4.3</td>
<td>Davy and Godsill [2002], Song et al. [2002]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neural Networks</td>
<td>Section 4.1</td>
<td>Augusteijn and Folkert [2002], Cun et al. [1990], Hazel [2000], Moya et al. [1993], Singh and Markou [2004]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering</td>
<td>Section 6</td>
<td>Scarth et al. [1995]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nearest Neighbor based</td>
<td>Section 5</td>
<td>Pokrajac et al. [2007], Byers and Raftery [1998]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Techniques</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table X. Examples of anomaly detection techniques used in image processing domain.
<table>
<thead>
<tr>
<th>Techniques</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification Based</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Clustering Based</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Nearest Neighbor Based</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Statistical</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Information Theoretic</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Spectral</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Applications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cyber-Intrusion Detection</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Fraud Detection</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Medical Anomaly Detection</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Industrial Damage Detection</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Image Processing</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Textual Anomaly Detection</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Sensor Networks</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table I. Comparison of our survey to other related survey articles. 1 - Our survey 2 - Hodge and Austin [2004], 3 - Agyemang et al. [2006], 4 - Markou and Singh [2003a], 5 - Markou and Singh [2003b], 6 - Patcha and Park [2007], 7 - Beckman and Cook [1983], 8 - Bakar et al [2006]
Lit Review Conclusion

- Your own thoughts
  - Evaluate the state-of-the-art
  - Promising directions
  - Open challenges
  - Broader relationship with entire discipline

- Presumably, you’re interested in the area of your lit review, so you should have thoughts!
  - Note that your own thoughts should be permeating the lit review as well, but the conclusion is a key place
Ph.D. Portfolio Lit Review

- This is a stand-alone lit review
  - You don’t have to talk about your own work in it at all
    - You may hint at promising/fertile research directions of course, which may be your own
  - It is NOT a Ph.D. Proposal!

- In fact, it does not have to be in your thesis area
  - Of course, it makes your life easier if it is since you’ll already have done the lit review

- You should start thinking about it early on!
How to make it Interesting!

- A big problem with lit reviews is that they are just boring to read
  - Especially those written for the PhD Portfolio (in general, not all!)
- What makes a lit review interesting is Critical Thinking
- It’s a “review” not a “summary”
  - Shouldn’t be 95% description of previous work and only 5% discussion
The Worst Lit Review

- Author 1 did something
- Author 2 did something else
- . . .
- Author n did something else
- Conclusion: a lot of work has been done and it’s all very impressive
How to make it interesting!

- Critical appraisal: evaluate strengths and flaws/limitations of reviewed works
  - While remaining civil
- Establish relationships between the reviewed works
  - Mention competing approaches/authors
  - Mention approaches subsumed by others
  - Pit approaches against each other based on published results/comparative studies
  - Mention when some comparative studies should be done but haven’t been done
How to make it interesting

- Identify open questions
  - From “future work” sections
  - From your own ideas

- Give a sense of historical development in the field
  - Identifying main phases of advances, so that you can tell the “story” of the field

- “Call” authors on their claims for novelty

- For each reviewed paper, knowing what they claim to accomplish and what they “punt on” is more important than how they do it in terms of structuring the lit review

- Write the lit review without giving any technical details and see if it reads well/interesting
  - Can you actually tell the story of the lit review in 10 minutes?

- Pick the topic appropriately
  - Writing a fascinating lit review is easier for some topics
The Not-Worst Lit Review

- The problem is this
- In one of the first works, Author 1 did this
  - It was great, but had one big problem
- Authors 2, 3, 4 tried to solve it, but unsatisfactorily for these reasons
- Then Authors 5 finally proposed a good solution launching a new era, that unleashed a whole set of new works
- Building on the work by Author 5, Authors 6 and 7 have proposed different approaches, and it's difficult to tell which one is best
  - Both claim greatness, but ...
- ....
- At this point, open questions are ...., and perhaps works by Authors 20, and 21 provide first steps toward answers.
Read Lit Reviews!

- You should be reading lit reviews
  - To learn about important topics
  - To understand what there is to do in a potential research area

- Where to find good lit reviews?
  - ACM Computing surveys is a standard source
  - Keywords like “survey”, “review”, “taxonomy”, help locate lit reviews
  - Lit reviews are often cited as well
Conclusion

- Writing a good lit review is known to be difficult
  - I review many bad ones for journals!
- It’s normal to go through many revisions of it as you write it
- But in the end the result is extremely useful, to you, and to others
- Read papers, read papers, read papers