Literature Reviews

Justin Zobel

University of Melbourne, Australia

Semester 2, 2018

Overview

Finding papers (and other materials).

Scoping the search.

Analysis of papers.

Towards a literature review.

Contents

Searching the literature

Analysis of papers

Literature reviews

Research literature

What is the "scientific literature" that must be consulted while undertaking research?

- Documents that are accepted by the scientific community as a reliable source of knowledge.
- Typically: reputable, refereed, widely accessible, citable (that is, uniquely identifiable), durable.

There are many sources of scientific knowledge:

- Books, journal papers, conference papers, workshop papers, theses.
- Sometimes includes web pages (Wikipedia?), technical reports, manuscripts.

Does not include *primary sources* such as lab notebooks, responses to a survey, outputs from an experiment, material that is being quantitatively analysed, ...

Research literature ...

By the time your research is complete, you need to be confident that you have seen and understood all the work that has a significant connection to your research.

Your reading achieves several aims:

- Establishes to your satisfaction that your work is novel.
- Helps you understand current theory, discoveries, and debates.
- Helps you to identify new lines of questioning or investigation.
- Provides fresh perspectives on your work.

This reading will ultimately be summarized in the background sections and the discussions of related work in your thesis (and your other writing).

Critical thinking

A key aim of reading is to develop *critical thinking* skills.

Successful researchers (e.g., people who complete a research degree) must demonstrate their ability to objectively analyze the work and claims of others – and of themselves.

Initially, you may explicitly working through a set of standard 'critical thinking' questions: 'what in the paper is new?', 'could the experiments be reproduced?', . . .

More subtle questions may also be appropriate. For example, for older work, it could be relevant to ask whether it was correct (in the context of current knowledge) at time of publication, and whether the lessons are still of value.

With experience, you place each paper in a context of other work that you know, and assessing it on a range of characteristics. You will become alert to common mistakes and bogus claims.

Other publications

News articles, science magazines, etc., may alert you to the existence of reputable work, but are almost never worth citing.

Locating the papers you need to establish the scientific foundations of a piece of research involves consulting the research literature.

. . . but your own learning may be built on a wider literature, including popular science texts, newspaper articles, web forums, etc.

Patents may also be relevant.

- There are many standard patent search resources.
- Many patents are junk.
- Ignore the claims.

There is no need to be exhaustive in your attempts to discover such material.

Finding relevant research

The process of search and discovery of useful papers can be thought of as a form of learning:

- Initial steps consist of guessing search terms, blindly using search tools, following up promising papers, identifying which individuals are prominent in the area.
- Each paper refines your understanding of the terminology, helps indicate which papers are significant, suggests new angles to follow up, further defines your criteria for whether a paper is “in” or “out”.
- As your understanding deepens, the search becomes easier and more complete.

Finding *all* relevant work is hard; for example, exhaustive professional searches across the medical research literature can take months of full-time work.

Finding all *significant* work is a critical part of doing research.

Finding relevant research ...

Another metaphor: find the *gateways*, follow the *paths*.

Gateways: papers that you discover by searching, chance (e.g., in seminars), or word-of-mouth.

- ACM portal, Citeseer, IEEE DL, Google scholar, Google/Live/Yahoo!, University search services, ...
- Search adaptively, narrowing or broadening the search in an exploratory way. "What will turn up if I try ...?" Don't expect to use just a few simple queries.

Paths:

- Chains of citation and cited-by (see the ACM portal).
- Other papers by the same authors, same research group.
- Other papers in the same journal, same conference.
- Web pages of significant authors, groups, projects.
- In some disciplines: registers of experiments, etc.

Competitors

A thorough literature search is likely to uncover a small number of other groups working in the same direction.

Competitors: Those whose activities mean that they may publish results that supersede or invalidate your own work.

(If there are too many competitors, and you don't have a strong competitive advantage, you may want to refocus your efforts onto a less popular problem.)

- Identify your competitors.
- Estimate how active they are, how many people are involved at each centre.
- Closely monitor their work.
- Take opportunities to establish dialogue with them.

Scoping the search

Scope: the field of papers that potentially have impact on your work.

It is rarely easy to identify the scope.

- Papers in unrelated areas can be highly relevant – consider the many different applications of caching.
- You may need to understand a paper to decide that it is unrelated to your work.
- The volume of literature may force you to narrow your inquiry.
- A large body of older work may be irrelevant in the light of new developments, or may be adequately encapsulated in a single text or survey paper.
- Or a critical older paper may have been neglected.

Scoping the search ...

A good strategy is to first uncritically *gather* and then later critically *analyze* and *categorize*.

Plan your search:

- Identify specific areas to be pursued.
- Set an area aside if the search appears to be exhausted.
- Revisit the key papers that competent authors are likely to cite – check for new citings.

Your topic and interests are likely to shift, focus, or broaden during your research – update your searching as you go.

A definite, final scope is only likely to be obvious once you have completed the investigative phase of your research *and* you have a good draft of your background chapter.

Contents

Searching the literature

Analysis of papers

Literature reviews

Initial reading

Identify the key result or idea in the paper (look beyond the abstract), asking questions such as:

- What are the researchers trying to find out?
- Why is it an interesting or useful problem?
- How original is the approach?
- What things were proposed?
- What things were measured?
- What were the results, that is, what do the authors conclude and what is the evidence?

A key exercise is to decide where the paper belongs.

- What other papers should it be classed with?
- Are the results current or superseded by newer work?

Initial reading ...

Take the time to absorb each paper. Browse it, get the general idea, revisit it until you feel that you have learnt enough.

There's no need to read a paper from end to end, or in order. Perhaps technicalities can be neglected; perhaps it will be most helpful to read "Related Work" first.

Set aside a paper if it is clearly not relevant, but don't rush to judgement.

(If the paper is heart-sinkingly similar to your own work,

- Use your expertise in the area to read it carefully.
- Maybe the work is excellent, but maybe it isn't.
- Maybe they have done groundwork for you, for example by collecting data or establishing the validity of the problem.
- There is almost always an angle they will have overlooked.)

Relevance

Papers will vary in their degree of relevance to your work. Assess them against criteria such as:

- Totally irrelevant, no connection.
- Tangentially relevant, throws light on related issues.
- Not actually relevant, but should be referenced because of superficial similarity.
- Highly relevant, a competing method whose shortcomings and strengths must be clearly identified.
- Foundational work on which your new results build.

(This list is not exhaustive.)

Papers can be relevant in indirect ways. For example, a paper may be useful because it proposes a methodology for a similar kind of experiment.

Correctness

A form of evaluation of a paper is to review it – the process of assessing whether it is suitable for publication.
(Reviewing is the topic of another lecture.)

But even a strong paper can be flawed; for example, most experimental papers have some aspects in which they could be improved.

The question is: to what extent do you accept the results?

- How were the claims evaluated or validated?
- What are the questionmarks over the experiments?
- Were appropriate baselines used?
- What are the limitations of the work?
- What conclusions do you draw from the data? Are the authors' claims reasonable or inflated?
- How clear and complete are the explanations?

Significance

Some papers describe results and inventions that are breakthroughs of genuine novelty; but most concern increments, of greater or lesser importance.

Papers of high impact need to be thoroughly understood. Other papers may – but may not! – require careful attention.

In assessing significance, also consider:

- Reputation of the institute, group, author.
- Age of the work. (Is a 1983 paper on software development practices likely to still be of value?)
- Accessibility and status of the publication venue.
- Reputation of the people associated with the venue (editorial board, program committee).

Other outcomes

A well-written paper can serve as a model for your thesis.

- Overall structure and style of presentation.
- Choice of elements to include and exclude.
- Methods of describing algorithms, experiments, results.

Badly written papers can provide useful negative examples. That is, take the effort to discover and understand the respects in which other papers are poor.

- A useful exercise is to explicitly review papers to identify shortcomings – for example, to ask why a result is unpersuasive.

Read with caution

A research paper is not a textbook:

- New ideas, not always well-digested, not always well-written.
- Detail is not always important.
- Some content may be incorrect.
- Each paper is an example of how to present results, do research, and make mistakes.

Minimize the time spent with each paper without compromising your learning.

Be critical, skeptical – evaluate the contribution, expect to be convinced, seek counter-arguments.

Read with caution ...

Some papers simply don't make sense. Uninteresting or wrong-headed ideas, badly tested, badly reported, inadequately reviewed. Common issues:

- Meaningless problems or topics.
- Arcane or pointless mathematics.
- Proof of the wrong property (such as complexity instead of correctness).
- Invention of arbitrary metrics.
- Unreasonable or implausible assumptions.
- Proof by assertion.
- Inflated claims and conclusions.

Learn from other researchers' mistakes.

Caveat: don't assume work is rubbish just because you don't understand it.

Read with caution ...

Promotional literature is not trustworthy.

Information on the internet (web forums, `slash.dot`, news groups) can represent a broad consensus.

. . . or it can represent a couple of opinionated people flaming each other.

Seek authoritative sources.

Anonymous information is almost always valueless.

Contents

Searching the literature

Analysis of papers

Literature reviews

Purpose of a literature review

Literature reviews are used to educate the reader – to introduce the knowledge that is needed to understand the contribution of the thesis.

They summarise the state of the art, and provide a critical analysis of previous work.

- Consolidation of diverse publications into a single discussion.
- Explanation of separate works in a common language and format.
- Organisation of the literature into themes.

Some papers are noted as relevant, and in some cases explained in a sentence, but without detail.

Significant papers are discussed as part of a flowing narrative that links them together.

Organisation

It can be helpful to draw a visual map of all of the references that you wish to discuss.

- Group together related papers.
- Describe groups with topic labels.
- Connect similar papers with lines.

The sections of the literature review should correspond to natural groupings.

Individual papers

Discussion of a single paper typically has the following elements.

- Identification in the text of the authors. ('Dent and Warborough [13] propose an alternative.')
- A summary of the contribution of the paper.
- A summary of the evaluation.
- Insights into the scope of the work and the validity or impact of the evaluation, that is, critical analysis.

These create understanding of each paper's contributions, limitations, and strengths.

The limitations of past work identified through the literature review provide the argument for undertaking new work. That is, they are the rationale for your own investigation.

Managing research literature

Investigating the literature in the context of a new (unfamiliar) area is likely to involve:

- Dozens of queries to each of the main search services.
- Browsing many hundreds, maybe thousands, of paper titles and abstracts.
- Emails to authors to obtain hard-to-get publications.
- Reading fifty to a hundred references to get background for a paper, several hundred or more to get background for a thesis.

You will need to make a strong presentation of this literature at the end of the research. This involves careful management of your knowledge of the literature along the way.

Maintaining a bibliography

- Keep a softcopy of every paper that was worth considering, even if it was ultimately deemed irrelevant.
- Organise the papers into folders as you process them – collections of PDFs are not easily browsed.
- Capture complete bibliographic data straight away:
  - Title, authors, venue, etc.
  - The URL the paper was obtained from.
  - Include a few keywords; for example, name the research group it came from and note the key contribution.
- Capture other data as soon as the paper is read:
  - Reflections on the correctness, etc., as discussed above.
  - Note any criticisms or features that may be of value later.

Start with a sensible format: use Bibtex or Endnote and a systematic method for forming identifiers.

Maintaining a bibliography ...

- Begin a document that cites your references.
- Use a reader-friendly citation style (e.g., Harvard).
- Group the citations by topic; papers should be listed together if you expect to discuss them together.
- Get into the habit of writing a sentence or two summarizing each paper; include experimental methodology as well as proposals and outcomes.
- Write another sentence or two of critical analysis: your view of strengths, limitations, and weaknesses.
  - How could the investigation have been improved?
  - Were all reasonable alternatives investigated?
  - Are there any obvious counter-arguments to the claims?

This document will evolve into your literature review.

Reading is an ongoing commitment

Expect to browse for new papers every couple of weeks, or more often.

Don't allow reading to develop into a deflection or procrastination activity – it needs to be part of a productive cycle of work, not a dominant use of time.

Expect to have a range of modes of reading:

- Browsing to find papers and get an overview of activity.
- Background reading of texts and popular sciences.
- Light or superficial reading of a broad selection of papers to understand the main outcomes in a body of work.
- Thorough, focused reading of key or complex papers that stretch your abilities or at the limits of your understanding; expect to need to use other resources to make sense of these papers.

Reading needs to be progressively documented.

Image courtesy Anna Zobel