

CS 2001: Research Topics
Reading Papers

Prof. Adriana Kovashka
University of Pittsburgh
September 8, 2021

Plan for this lecture

- Structure of a paper
- Reading papers for breadth/depth
 - How to maximize time efficiency
 - Three-pass approach
 - How to organize papers
- Homework (activities)
 - Read 1 iconic paper (questions at high/mid level)
 - Read 4-6 iconic papers (compare/contrast)
 - Google Scholar search for a topic of interest

Structure of a Paper

Structure of an 8-page, 2-col Paper

- Abstract
- Introduction
- Related Work
- Approach/Method
- Validation/Results/Experiments
- Conclusion
- Acknowledgements
- References

Abstract (1 paragraph)

- Purpose
 - Why is this paper interesting?
 - Why should I spend my time reading this?
 - What do you claim to do?
- Summarize problem being tackled, key innovations, main results
- One strategy: write it after rest of paper written
 - Caveat: some conferences might want an abstract earlier than the submission deadline

Introduction (1-1.5 page)

- Purpose: motivate work, inform reader what is to come
- First paragraph: describe problem and state current gap in science, including current methods' deficiencies
 - Also briefly describe motivation: why addressing this particular gap is important, has applications
- Second paragraph: state the very high-level idea of your solution to address this gap
- Third paragraph: briefly introduce your method
- Fourth paragraph: briefly summarize experiments
- Fifth paragraph: summarize contributions
- Concept figure (Fig. 1): illustrate key idea of the work
- Important: sets initial expectations

Related Work (0.5-1 page)

- Note: This may also come before Conclusions—tradeoffs?
- Purpose: inform reader you are aware of prior results, and demonstrate the novelty of the work
- Summarize 15-50 most relevant papers
- Organize by rough topic or approach (use paragraph headings but no separate subsections, for space)
- Within each paragraph, can describe papers at a course level, e.g. “Some works approach the task by... [ref1, ref2, ref3].”
- End each paragraph or group of papers with how proposed method is different

Approach (2.5-3.5 pages)

- Purpose: describe novel approach, contribution
- Have an introductory paragraph describing an overview of the approach
- Make it clear which part of the approach is standard/from prior work, and which part is new/novel/innovative
- Common to have 2-5 subsections
- Should be detailed enough for others to replicate the work (in theory)—in ML, common to include a separate “implementation details” section

Experiments (2.5-3.5 pages)

- Purpose: prove stated contributions are meaningful
- Forms: performance measurements, simulation results, analysis of user study data, formal proofs...
- Include experimental setup (e.g. data and metrics used, methods tested with names, perhaps in special font setting)
- Include one or more results subsections (the outcomes from running the experiments)
- Show results in tables and/or figures, and describe/analyze these tables/figures in the text
- Explicitly make all the conclusions you want the reviewers to draw from them
- May or may not show statistical significance but be prepared to answer questions about it

Conclusion (1 paragraph)

- Summarize contributions and findings of the paper, briefly
- Acknowledge limitations and briefly suggest future work– no more than 1-2 sentences (may vary by research community)

Discussion

- May be part of experiments, potentially conclusion
- Purpose: papers do not often “close” a topic—this is where you reflect on what has been done, and what is still open
- May include: interpretation of results or evaluation, discussion of open problems, description of limitations, etc.

Acknowledgements (1 paragraph)

- Omit for review due to anonymity
- Ask your advisor which funding sources to acknowledge, any particular phrasing to use
- Can also thank the anonymous reviewers and area chairs

References

- Standard formats for this
- Can get bibtex citations for a paper from Google Scholar or DBLP, but may have to fix
- If a paper appeared in a conference/journal, don't cite the arxiv version
- Use consistent formatting, e.g. the same conference should use the exact same string including abbreviation for the conference name

How to Read Papers

Why do we read papers?

- To know what others in our field are doing
- To get ideas for research projects
 - in terms of problems/tasks, as well as gaps and deficiencies in prior methods
- To know the relevant techniques
- To avoid reinventing the wheel
- To avoid having our papers rejected because we didn't compare to relevant baselines

Why do we read papers?

There are many legitimate reasons for reading a paper

- I heard someone talking about this result
- The conference talk interested me
- It's related to a problem I am working on
- My advisor told me to
- This provides context for another problem
- I think that I might want to explore this area
- ...

Curiosity

Depth

Breadth

Take-away point: *Why* you plan to read a paper will—to some degree—dictate *how* you should go about reading it

Too many papers!

- CVPR, the top conference in my field, has over 1000 papers each year— and there are 5+ such conferences each year in my field
- Sometimes papers are organized by topic
- Last CVPR, I went through all titles and downloaded 300 papers, read the abstract and skimmed intro and figures
- Took notes on about half of them—otherwise I'd forget which papers did what
- Faculty often skim papers, and you will skim some, but have to read some in detail

Reading for breadth vs depth

- Breadth
 - Get a sense of the “pulse” of the field, what are open problems
 - Identify potential research directions
 - Get exposed to techniques—you never know when a seemingly unrelated technique may come in handy
- Depth
 - If your advisor sends you a paper, spend at least 30 min reading
 - You need to know papers in great detail, your advisor less so—the details are your responsibility
 - But if you do 5 hours of reading a week, you should spend 20 hours doing active work, i.e. developing your method and conducting your experiments

Maximizing Time Efficiency: Example

- Suppose you have 100 papers roughly relevant to your area
- Read abstracts+intros for all, look at concept/method figures (5-10 min per paper)
- Also skim approach and experiments for 20 (30 min per paper)
- Read 20 in great depth (2-4 hours per paper)

Multi-pass Approach

Keshav* has a nice paper on a three-pass reading approach

Pass 1: Basic comprehension

- What is the main topic of the paper?
- What are the authors' claimed contributions?
- What do they cite?

Pass 2: First look at real details

- Focus on details: evaluation, figures, methods
- Ignore proofs

Pass 3: Depth!

- Fully understand all details

* S. Keshav, "How to Read a Paper," ACM SIGCOMM Computer Communication Review 37(3) : 83-84, July 2007.

First Pass: Decide How Much Time to Invest

Focus your attention on:

- Title and Abstract
- Full details of the Introduction
- Section and Sub-Section headings in the body
- Full details in the Conclusion
- Skim references, note what you've read

After this, you should know about the “5 Cs”

- **Category**: Experimental paper? Theory? Measurement?
- **Context**: What does this paper cite?
- **Correctness**: Do any assumptions seem reasonable?
- **Contributions**: What do the authors (claim) to contribute?
- **Clarity**: Can you follow the paper?

You can probably accomplish this for most papers in ~10 minutes

Activity 1 (10+10 min)

- Pick a paper from the iconic papers list
- Answer 5C's (take notes)
- Share with class

Second Pass: Breadth of Knowledge

General idea: Read the whole paper, but skip super-intricate details like proofs, some equations (if many).

Focus on:

- Understanding methodology, evaluation, figures, etc.
- Mark relevant references for later reading (**more breadth!**)
- Being able to explain the main ideas of the paper to someone else

This process can take **up to an hour** for a 10-page paper

Why so long?

- Perhaps you're new to the subject area
- Authors use methodologies or techniques that are unfamiliar
- Paper is just badly written...

Second Pass: Breadth of Knowledge

Important questions include:

- What are the **motivations** for this work?
- What is the proposed **solution**? Is it **novel**?
- How is this solution **evaluated**?
- **What do you think** about the problem, solution, and evaluation?
- What are the **contributions** of this work?
- Does this paper **close** an area of research? **Open** a new one? Lead to interesting **future work**?
- What **questions** do you still have?

Griswold has a nice template for answering these questions—linked on the course page.

Activity 2 (60 min, HW)

- Pick another paper from the iconic papers list
- Answer questions from previous slide (take notes)
- Share with class next time

Third Pass: Develop Intricate Understanding

Main focus: Everything you've glossed over so far

- Thorough scrutiny of assumptions
- What alternative solutions might have been possible?
- Does the evaluation cover enough meaningful cases?
- Detailed examination of proofs and proof techniques

After a thorough pass, you should (ideally) be able to replicate the results presented in the paper

This is a **time-intensive** process

- 4-5 hours for beginners
- Around an hour for more experienced readers

Note-Taking

Note-taking can help build your understanding of a paper and manage the *many* papers that you'll eventually read

Note taking **while you read** helps capture the context of your reading session for later reference

Use a highlighter to mark major points, definitions, and theorems for quick reference later

Make notes in the margin

- Write down questions as they pop into your head
- Answer previous questions as you find answers
- Summarize tables, graphs, etc.
- Add details to incomplete/unclear examples

Note-Taking

Note taking **after you read** can also help

- Ensure complete understanding of relevant papers
- Manage large collections of papers as your progress in your studies

Consider making a document per research area

For each paper, write up:

- A technical summary of the work
- A brief description of the paper's relation to other works
- Relationships to your ongoing/planned research
- Resources you can use in your work (e.g. data)
- Any cool ideas for future work that come to mind

Filling in the Gaps

Initially, you will have an incomplete knowledge of a research area. How can you fix this problem?

Step 1: Read up on prior work!

Step 2: Understand how this paper fits into more recent research

There are research tools to help aid these processes

- ACM portal: <http://portal.acm.org>
- IEEEXplore: <http://ieeexplore.ieee.org/Xplore>
- Google scholar: <http://scholar.google.com>
- Citeseer: <http://citeseerx.ist.psu.edu/>
- DBLP: <https://dblp.uni-trier.de/>

How to Organize Papers

- I used to keep folders with printed and annotated papers, labeled by topic/subarea
- I still annotate papers, and keep a separate document with notes (for easier searching)
- Reference management software support exists:
 - Mendeley
 - EndNote
 - Zotero
 - ReadCube Papers

Activity 3 (HW)

- Read 2-3 iconic papers in one area, and 2-3 in another
- Compare and contrast: Summarize and present writing strategies– how papers in this area are typically written
- Highlight any differences between areas, any similarities (share)

Activity 4 (HW)

- Pick a topic or a research project idea (should be fairly specific, not a whole research area)
- Do a Google Scholar search to determine what others have worked on in this area
- Share with us what you found

Activities 2, 3, 4