

Topic: How to read a Research Paper

Introduction

Researchers spend a great deal of time reading research papers. Why they are reading research papers? There are many different reasons, e.g.:

- For a literature survey of a new field/problem
- Be up-to-date on current research in the field
- Allows you to replicate/extend the results
- Provides you with useful data
- Gives you “pre-digested” thoughts
- To decide whether (and where) to publish
- Teaches you how to write
- Review for a conference or a class, etc.

Learning to efficiently read a paper is a critical but rarely taught skill. Beginning undergraduate students, therefore, must learn on their own using trial and error. Students waste much effort in the process and are frequently driven to frustration.

Today, we will study the special approach which can help to improve the paper reading process.

1. Types of (Computer Science) Research Papers

For clarity let's enumerate the main kinds of research papers:

- Conference Papers – most recent, “hot of the presses”, information.
- Technical Reports – expand on the information in a conference paper.
- Journal Papers (or Articles) – expand and combine results from several conference paper.
- Book Chapters – expand conference paper or journal article.
- PhD Dissertations – frequently revised into either journal papers or book chapters.

- Posters – extended abstracts, short, good for discussions.
- Workshop Papers – very hot / new topics, work in-progress, ideas, preliminary results.

NOTICE: Conference papers and journal papers are “peer-reviewed” have been examined by other computer scientists (2-5) prior to publication (“double-blind” review). Technical reports are typically not peer-reviewed, but are still excellent sources of detailed information. Posters and workshop papers are peer-reviewed (less number of reviewers).

2. Active reading

First of all, before you start to read anything, think about why you are doing so. This is quite important as your intentions determine the effectiveness of your reading, so your needs influence how you will read an article, book – literally anything. However, if you just need to have an overview in order to get an idea of what you can use for your own research a quick skim-through will suffice. But, if you are planning to present certain findings from journal articles you might need to dig into it more deeply. Take into account that maximizing what you read isn't just about adjusting the surrounding environment that allows you to concentrate better. You need to learn how to react to what you read to become a mindful and active reader!

Active reading includes these steps:

1. Get rid of distractions.
2. Get a pen.
3. Jump around, re-read, go backwards. – These are not novels!
4. Talk to others.

When you read, you have to keep in mind the several questions:

- What questions does the paper?
- Is the problem relevant?
- What are the main conclusions of the paper?
- What evidence supports those conclusions?
- Do the data actually support the conclusions?
- What is the quality of the evidence?
- Why are the conclusions important?

However, we also need an effective reading process. It can be a multi-pass reading. The main idea centers on going through many passes that reach the purpose of your reading. By using this method, the common approach that one should read a paper from top-to-bottom is not carried out here. Specifically, the first pass gives a reader a general idea about the paper and the second one helps to grasp its content. Then third pass allows one to understand an article in more depth, etc. This approach is not only helpful for estimating the time required to reading a set of papers, but also for deciding on how much in-depth you want to research the paper.

Thus, thanks to implementing this approach one can become a mindful and active reader by spending less effort and time on one's readings. Sounds good, doesn't it? Now, let us look on three-pass method from Keshav [1]. We'll go through each pass carefully in order to have a quick glance at the process itself.

3. Three-pass reading

Keshav propose to use three pass only. Each pass accomplishes specific goals and builds upon the previous pass: The first pass gives you **a general idea** about the paper. The second pass lets you grasp **the paper's basic content, but not its details**. The third pass helps you **understand the paper in depth**.

Let's focus on [the first pass](#) at the moment. It centers on getting a bird's-eye view of the read paper so performing a quick scan that takes up to **5-10 minutes**. Moreover, it helps to determine whether the paper will be useful or of interest and make it easier for you to validate some of the author(s)' claims. At the same time, simply skimming through the paper allows you to gauge the level of knowledge you require to fully understand it. It is worth noting that reviewers use this method quite often to decide whether they should accept or decline a manuscript in a review process. That is why, it is commonly emphasized that one should pay extra attention on preparing a clear, concise and comprehensive abstract as well as choosing good titles for sections and subsections in one's manuscript. With this in mind, let's follow the five simple steps of the first pass:

1. Read the **title**, **abstract**, and **introduction** with care.
2. Ignore everything apart from the **section and sub-section headings**.
3. Go through all the available **mathematical / statistical content** to determine the underlying theoretical foundations of the paper.
4. Go through the **conclusions**.
5. Browse over the **references** and mentally tick the ones you've read already.

Now try to answer five C's when you are done.

- Category: What type of paper is this? Is it an experimental study? A combined analysis of previous studies? An introduction of a new methodology?
- Context: Which other literature papers is it related to? What were the theoretical bases that were used to analyze the problem?
- Correctness: Are the assumptions valid?
- Contributions: How does the paper contribute to the existing scientific understanding?
- Clarity: Is the paper well-written and easy to understand?

Now that we have gone through the first pass, let's have a look at [the second one](#). Here, we dig into the paper more in-depth but omit details such as proofs. It is helpful to mentally summarize the text into **key points** and make **notes** as we read (*look on next chapter too*). But do not forget to write down terms you do not fully understand or some inquiries you would want to ask the author himself. At this stage, it is important to:

1. Examine presented figures, diagrams and other illustrations, by paying special attention to graphs. Answer these questions:
 - Are the axes properly labeled?
 - Are the results shown with error bars, so that conclusions are statistically significant?
 - This will help you to distinguish good work from those that are poorly done.
2. Do not forget to **mark references** you find relevant. It is important for later readings and a great way to learn the backgrounds of the studied paper.

Now, we should be able to summarize the paper into a few short key points along with some supporting evidence. The depth of reading during the second pass is appropriate for someone who may or may not be familiar with the field of the paper in question, but is interested in a paper to some extent and finds it useful. This process can take up to **one hour** of work and requires a moderately high level of concentration.

You can now choose to:

- set the paper aside, hoping you don't need to understand the material to be successful in your career,
- return to the paper later, perhaps after reading background material or
- persevere and go on to the third pass.

The **third and final pass** is helpful for readers who need to **fully apprehend** the contents of the paper, as is in the case of manuscript reviewers. What is the most important here is to make an effort to “**virtually re-implement**” the paper – that means to re-create the work and assumptions of the authors. This pass requires a **huge attention to details** as innovations and assumptions. Moreover, some hidden authors’ failures may be identified here. Through the process of virtually re-implementing the work you challenge your thoughts and redirect them to approach the paper proactively. Here, you are also focusing your attention on writing down possible **ideas or recommendations for future research**. The third pass can take **from three to five hours** for beginners and up to **one hour** for an experienced reader.

At the End of this Pass you should be able

- to **reconstruct** the entire structure of the paper from memory;
- to identify its **strong and weak points**;
- to pinpoint **implicit assumptions**, **missing citations** to relevant work, and **potential issues** with experimental or analytical techniques.

4. Notes writing

Write as you read. This keeps your attention focused and makes you engage with the paper.

Low-level notes

Often it is easiest to scribble notes on the printed-out paper itself, responding in context to the formulas, figures, and text. In that case, file or scan your annotated copy for future reference (or perhaps annotate the e-document directly, without printing or scanning).

You can use notes on the paper to

- restate unclear points in your own words;
- fill in missing details (assumptions, algebraic steps, proofs, pseudocode);
- annotate mathematical objects with their types;

- come up with examples that illustrate the author's ideas, and examples that would be problematic for the author;
- draw connections to other methods and problems you know about;
- ask questions about things that aren't stated or that don't make sense;
- challenge the paper's claims or methods;
- dream up followup work that you (or someone) should do.

High-level notes

Low-level notes aren't enough. Also keep high-level notes about papers. It's quite useful to distill the paper down: summarize the things that interested you, contrast with other papers, and record your own questions and ideas for future work. Writing this distillation gives you a goal while reading the paper, and the notes will be useful to you later.

Michael Mitzenmacher writes: “Read creatively. Reading a paper critically is easy, in that it is always easier to tear something down than to build it up. Reading creatively involves harder, more positive thinking. What are the good ideas in this paper? Do these ideas have other applications or extensions that the authors might not have thought of? Can they be generalized further? Are there possible improvements that might make important practical differences? If you were going to start doing research from this paper, what would be the next thing you would do?”

I suggest sorting your file of notes chronologically, by when you read the paper, since that may help you find vaguely remembered papers or remember what else you were reading at the time. Sometimes you'll want to search by author/title/etc., so start the notes for each paper with a rough citation.

5. Remembering Read-Papers

Strongly recommend that you make an electronic file for your own bibliography. After reading, if you think it's worth remembering, write an entry for that paper in your bibliography file.

You should note: authors' names, paper title, venue, date of publication, and page numbers (if possible).

Add a 2-3 sentence description of the paper in which you summarize for yourself the problem addressed by the paper, the solution proposed or results learned, and main contributions.

Conclusion: Literature Survey

What should be noted is that the multi-pass approach can also be considered as a really helpful method to use when doing literature research. Usually, we focus our search in the field of our specialization or interest, but not always. The facts are that in order to find the right papers for our projects, we need to dig through tons of research articles. How do we go about using this approach in literature research?

1. Search using an academic search engine (e.g. Google Scholar, MS Academic Search, CiteSeer, or ACM digital library) by typing in specific keywords you have found from three to five recent papers on the topic. How do you find the keywords? Use the first pass here – get the gist of the paper, then go to the list of referenced papers presenting latest works in the field. If they direct you to recent literature paper you can just search for it, read it, and congratulate yourself for being lucky.
2. If you have no access to such a paper, look for repeating author names and shared citations in the text (usually the main researchers or papers in the area). Find their articles, set them aside and go to the author's websites to see what they have published recently or where they have done so. These steps are useful not only for acquiring needed data, but also for identifying top conferences in the field.

3. Next, go to the conferences websites and browse through their latest proceedings which will allow you to identify latest, high-quality work in the field. Do a first pass and identify the most relevant works for you now. Then, along with those papers which have been set aside previously, perform the first and the second pass now.

Thus, as we can see, this approach is not only useful to help us out in becoming mindful and effective readers, but also when doing literature research.

Self-study materials

- 1) S. Keshav. How to Read a Paper // ACM SIGCOMM Computer Communication Review, 2007, Volume 37, Number 3, pp. 83-84.
URL: <http://ccr.sigcomm.org/online/files/p83-keshavA.pdf>.
Abstract: This article outlines a practical and efficient three-pass method for reading research papers.
- 2) Philip W. L. Fong. How to Read a CS Research Paper? 2004. URL: <http://www2.cs.uregina.ca/~pwlfong/CS499/reading-paper.pdf>.
Abstract: This article highlights some points a young researcher should bear in mind when reading a Computer Science research paper.