

Reviewing Papers

Three-pass approach

From Keshav “How to Read a Paper”

The key idea is that you should read the paper in up to three passes, instead of starting at the beginning and plowing your way to the end. 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 content, but not its details.

The third pass helps you understand the paper in depth

First-pass

1. Carefully read the title, abstract, and introduction
2. Read the section and sub-section headings, but ignore everything else
3. Read the conclusions

5Cs

1. Category: What type of paper is this? A measurement paper? An analysis of an existing system? A description of a research prototype?
2. Context: Which other papers is it related to? Which theoretical bases were used to analyze the problem?
3. Correctness: Do the assumptions appear to be valid?
4. Contributions: What are the paper's main contributions?
5. Clarity: Is the paper well written?

Second-pass

1. Look carefully at the figures, diagrams and other illustrations in the paper. Pay special attention to graphs. Are the axes properly labeled? Are results shown with error bars, so that conclusions are statistically significant? Common mistakes like these will separate rushed, shoddy work from the truly excellent.
2. Remember to mark relevant unread references for further reading (this is a good way to learn more about the background of the paper).

Second pass

After the second pass, you should be able to grasp the content of the paper.

You should be able to summarize the main thrust of the paper, with supporting evidence, to someone else. This level of detail is appropriate for a paper in which you are interested, but does not lie in your research speciality.

You can now choose to: (a) set the paper aside, hoping you don't need to understand the material to be successful in your career, (b) return to the paper later, perhaps after reading background material or (c) persevere and go on to the third pass

Third-pass

The key to the third pass is to attempt to virtually re-implement the paper: that is, making the same assumptions as the authors, re-create the work. By comparing this re-creation with the actual paper, you can easily identify not only a paper's innovations, but also its hidden failings and assumptions.

This pass requires great attention to detail. You should identify and challenge every assumption in every statement. Moreover, you should think about how you yourself would present a particular idea. This comparison of the actual with the virtual lends a sharp insight into the proof and presentation techniques in the paper and you can very likely add this to your repertoire of tools. During this pass, you should also jot down ideas for future work.

This pass can take about four or five hours for beginners, and about an 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, as well as be able to identify its strong and weak points. In particular, you should be able to pinpoint implicit assumptions, missing citations to relevant work, and potential issues with experimental or analytical techniques.

Types of papers

From Stent “How to Read a Computer Science Research Paper”

There are three **basic** types of CS research paper: theoretical, engineering and empirical.

A theoretical paper describes a theory or algorithm or provides a mathematical proof for some hypothesis.

An engineering paper describes an implementation of an algorithm, or part or all of a computer system or application. Engineering papers are now frequently required to include descriptions of system evaluation.

An empirical paper describes an experiment designed to test some hypothesis.

How do I know it is a good paper?

- 1) The problem the paper addresses is clearly stated, both in the abstract and early on in the paper itself. The technical importance and broader impacts of the paper are described.
- 2) The paper includes a clear description of the experiment, system or theory the problem addresses. This is usually the second section of the paper.
- 3) The paper describes and analyzes the results of the work described (either experimental or evaluation results).
- 4) The authors have some sound, non-trivial ideas for future work. This usually appears at the end of the paper.
- 5) Related work is described and cited correctly. You can get an idea of this by looking at the list of references at the end of the paper.

How to read, again?

First, skim the paper to check that it has the elements described above (clear problem statement, clear description of method or system, clear presentation of results, appropriate citation of related work). At the end of this quick skim you should be able to write down the type of the paper (theoretical, engineering or empirical), the area of computer science the paper is about, and the problem the paper addresses.

How to read, again?

Now read the introduction carefully, looking for the problem statement, theoretical importance and broader impact. Sometimes related work is also addressed in the introduction. Consider:

Are the author's claims reasonable and realistic?

Do the authors present evidence that they know why they are doing this piece of research?

Do they have an idea of the larger picture?

How to read, again?

Now read the methods section carefully (this section might be called “experiment” or “system description”, or may involve statement of theorems and their proofs). Ask yourself:

1. Can you think of counter-examples for examples given?
2. Is the approach clearly described? Can you outline the steps or summarize the approach?
3. Does the work address the problem stated earlier in the paper?
4. Does the approach seem to require unreasonable amounts of human guidance?
5. Does the approach seem objective?

How to read, again?

Computer science papers are often written in English by non-native speakers of English. Syntactic errors or awkwardness of phrasing do not indicate that the research is bad; you should try **distinguish between the writing style and the research** itself. If you get confused, you can reread or write down your questions to check in other work by the same authors.

How to read, again?

If, at the end of the methods section, you think the approach is sound (even if you disagree with it!), read the results, evaluation or analysis section. (If there is no such section and the paper dates from post-1995, there is a problem!)

Ask yourself:

- 1. Do these results address the problem stated at the start of the paper?**
- 2. Do the authors use commonly-accepted terminology and methods?**
- 3. What is the conclusion of the research?**

Finally, look at the conclusions and future work (last section of the paper). These can be good sources of topics for your own research. A brief note about older research papers: computer science is a rapidly evolving field and is still very young compared to many of the fields closest to it (mathematics, engineering). Research papers that date from before the 1990s may appear very different from newer papers. They may be more speculative, include less evaluation or no evaluation, and typically involve the use of much less data. There are very practical reasons for this (slower computers, less data available). These are not reasons to discount these papers!