How to Read a Paper

S. Keshav
David R. Cheriton School of Computer Science, University of Waterloo
Waterloo, ON, Canada
keshav@uwaterloo.ca

ABSTRACT

Researchers spend a great deal of time reading research papers. However, this skill is rarely taught, leading to much wasted effort. This article outlines a practical and efficient *three-pass method* for reading research papers. I also describe how to use this method to do a literature survey.

Categories and Subject Descriptors: A.1 [Introductory and Survey]

General Terms: Documentation.

Keywords: Paper, Reading, Hints.

1. INTRODUCTION

Researchers must read papers for several reasons: to review them for a conference or a class, to keep current in their field, or for a literature survey of a new field. A typical researcher will likely spend hundreds of hours every year reading papers.

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

For many years I have used a simple approach to efficiently read papers. This paper describes the 'three-pass' approach and its use in doing a literature survey.

2. THE THREE-PASS APPROACH

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.

2.1 The first pass

The first pass is a quick scan to get a bird's-eye view of the paper. You can also decide whether you need to do any more passes. This pass should take about five to ten minutes and consists of the following steps:

1. Carefully read the title, abstract, and introduction
2. Read the section and sub-section headings, but ignore everything else
3. Read the conclusions
4. Glance over the references, mentally ticking off the ones you've already read

At the end of the first pass, you should be able to answer the *five Cs*:

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?

Using this information, you may choose not to read further. This could be because the paper doesn't interest you, or you don't know enough about the area to understand the paper, or that the authors make invalid assumptions. The first pass is adequate for papers that aren't in your research area, but may someday prove relevant.

Incidentally, when you write a paper, you can expect most reviewers (and readers) to make only one pass over it. Take care to choose coherent section and sub-section titles and to write concise and comprehensive abstracts. If a reviewer cannot understand the gist after one pass, the paper will likely be rejected; if a reader cannot understand the highlights of the paper after five minutes, the paper will likely never be read.

2.2 The second pass

In the second pass, read the paper with greater care, but ignore details such as proofs. It helps to jot down the key points, or to make comments in the margins, as you read.

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).

The second pass should take up to an hour. After this 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.

Sometimes you won't understand a paper even at the end of the second pass. This may be because the subject matter is new to you, with unfamiliar terminology and acronyms. Or the authors may use a proof or experimental technique that you don't understand, so that the bulk of the paper is incomprehensible. The paper may be poorly written with unsubstantiated assertions and numerous forward references. Or it could just be that it's late at night and you're tired. 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.

2.3 The third pass

To fully understand a paper, particularly if you are reviewer, requires a 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.

3. DOING A LITERATURE SURVEY

Paper reading skills are put to the test in doing a literature survey. This will require you to read tens of papers, perhaps in an unfamiliar field. What papers should you read? Here is how you can use the three-pass approach to help.

First, use an academic search engine such as Google Scholar or CiteSeer and some well-chosen keywords to find three to five *recent* papers in the area. Do one pass on each paper to get a sense of the work, then read their related work sections. You will find a thumbnail summary of the recent work, and perhaps, if you are lucky, a pointer to a recent survey paper. If you can find such a survey, you are done. Read the survey, congratulating yourself on your good luck.

Otherwise, in the second step, find shared citations and repeated author names in the bibliography. These are the key papers and researchers in that area. Download the key papers and set them aside. Then go to the websites of the key researchers and see where they've published recently. That will help you identify the top conferences in that field because the best researchers usually publish in the top conferences.

The third step is to go to the website for these top conferences and look through their recent proceedings. A quick scan will usually identify recent high-quality related work. These papers, along with the ones you set aside earlier, constitute the first version of your survey. Make two passes through these papers. If they all cite a key paper that you did not find earlier, obtain and read it, iterating as necessary.

4. EXPERIENCE

I've used this approach for the last 15 years to read conference proceedings, write reviews, do background research, and to quickly review papers before a discussion. This disciplined approach prevents me from drowning in the details before getting a bird's-eye-view. It allows me to estimate the amount of time required to review a set of papers. Moreover, I can adjust the depth of paper evaluation depending on my needs and how much time I have.

5. RELATED WORK

If you are reading a paper to do a review, you should also read Timothy Roscoe's paper on "Writing reviews for systems conferences" [2]. If you're planning to write a technical paper, you should refer both to Henning Schulzrinne's comprehensive web site [3] and George Whitesides's excellent overview of the process [4]. Finally, Simon Peyton Jones has a website that covers the entire spectrum of research skills [1].

6. A REQUEST

I would like to make this a living document, updating it as I receive comments. Please take a moment to email me any comments or suggestions for improvement. You can also add comments at CCRo, the online edition of CCR [5].

7. ACKNOWLEDGMENTS

The first version of this document was drafted by my students: Hossein Falaki, Earl Oliver, and Sumair Ur Rahman. My thanks to them. I also benefited from Christophe Diot's perceptive comments and Nicole Keshav's eagle-eyed copy-editing.

This work was supported by grants from the National Science and Engineering Council of Canada, the Canada Research Chair Program, Nortel Networks, Microsoft, Intel Corporation, and Sprint Corporation.

8. REFERENCES

[1] S. Peyton Jones, "Research Skills," http://research.microsoft.com/ simonpj/Papers/giving-a-talk/giving-a-talk.htm.

[2] T. Roscoe, "Writing Reviews for Systems Conferences," http://people.inf.ethz.ch/troscoe/pubs/review-writing.pdf.

[3] H. Schulzrinne, "Writing Technical Articles," http://www.cs.columbia.edu/∼hgs/etc/writing-style.html.

[4] G.M. Whitesides, “Whitesides’ Group: Writing a Paper,” http://www.che.iitm.ac.in/misc/dd/writepaper.pdf.
[5] ACM SIGCOMM Computer Communication Review Online, http://www.sigcomm.org/ccr/drupal/.