

# Scientific Reading

Benjamin Bengfort

November 20, 2014

Managing the reading of scientific papers is a challenge that most graduate students face because, let's face it, scientific papers are tough to read. I would say that the first year of graduate school for me was an exercise not in learning how to do research or even higher order computer science principles but rather how to decipher academic papers (generally, not just in a single discipline). I'm still not sure that I'm an expert in reading these types of papers, but I am far more comfortable cracking open a paper and reading it than I was before; and importantly, I'm much faster at reading these papers than when I started.

Reading is the first step; comprehension is the second, especially when the papers are required to directly inform what you're thinking either for simple research papers, for coursework, or so that you can come up with new ideas. When I learned to read, I was trained in simple exercises in reading comprehension (read a passage, respond to questions). Unfortunately, narrative reading comprehension is not the same as scientific comprehension because critical evaluations of academic papers only partially use fact based extraction.

Consider if you were reading *Philosophiæ Naturalis Principia Mathematica* (one of the first scientific papers although only cited 1,750 times according to Google Scholar), comprehension questions such as "What is the inverse square law?" only reveal partial understanding of the text. Better questions would allow you to elucidate why the *Principia* is important, what work it evolved from, and what future work is required as a basis. For example, it's important because it demonstrates a new mathematical technique that we call calculus, and uses the calculus methodology to demonstrate a precise method for computing elliptical motion around the sun, thereby proving the existence of a gravitational force. A heliocentric model of the solar system, as well as work in optics, and the discovery of the inverse square law were crucial prior work. Finally, important future work involves discovering the medium through which gravity is transferred and why celestial bodies at great distances experience the instantaneous effects of gravity.

Frankly, that level of comprehension is provided to me through the benefit of nearly 400 years of reading comprehension with apologetic discourse. You don't get nearly the same thing with papers written in the last 7 years, reviewed most likely by graduate students who weren't willing or able to give a high level of scrutiny during the peer review process. You might also think that in computer science there would be software to review and execute; but this is also simply

not true. Very often code and data are not provided to repeat experiments in a meaningful way. We are left requiring a mechanism to critically evaluate text.

This semester I'm taking a class with Peter Keleher, and as part of the course we are reading papers related to the subject matter (distributed storage systems). Dr. Keleher is having us write blog posts on every paper as part of our course participation grade, and in these blog posts, we are asked to comment on or discuss the following:

1. What is the problem that the paper is attempting to address?
2. What is innovative about this paper, or what is the contribution?
3. Critically evaluate the paper with both positive and negative comments.

Over the course of the semester, I've discovered that this simple methodology of writing 500 words or so commenting on the paper in that fashion has led to a better reading comprehension on my part. Not only that, I've become faster at reading papers, because now I'm reading to answer these specific questions, but not only answer the questions - but comment on them and write about it in a coherent way. I'm finding that I'm doing a first pass read of the paper at a deeper than skimming level to get a feel for the organization, the topics, and any critical knowledge required. After that, I do a Google search for the topic, or review the bibliography for any linked knowledge I can add (which was tough when I first started, but now I'm able to recognize links to common systems). Finally- I start to write, and as I write, I dive back into the paper to read or understand critical concepts that are part of my elucidation of the three points above.

Needless to say, this seems like an extremely good habit to continue with as much reading as I can possibly get away with; and this blog seems like a perfect forum to publish my thoughts on other papers. And I'm hoping to make this a regular habit here. So stay tuned for more.

This methodology is addressing another problem I have as well: organization. Currently, I use Zotero and Papership to organize my research and reading. Once I get into a paper, I create a "reading list" of citations along with their abstracts; then some notes about their importance to the work I'm doing. Using the "blog post reading" methodology, I will have all of the pieces prior to the paper writing, pre-organized for my benefit. I believe that this process will be effective to making me a better scientist, and a better writer; and will update you as I go on! The first paper that I will experiment on will be a paper on diversity in Computer Science that has been sent around the department. It seems like an excellent starting place before getting into papers that are related to my research!