Computer Science and Engineering: The Bridge Between Disciplines

Jesus Gonzalez Castellon
Kelvin Lam
Rolly Padua
Majok Ring
Huimin Zhang

Abstract

The computer science and engineering discipline is unique in the lack of prominence of both survey and experimentation based research, as well as its simple information dispersal methods. While other fields of study may prioritize academic journal entries and publications, computer science sees its most prestigious research articles published through conferences, and welcomes less stringent restrictions on formatting.

Although computer science and engineering research remains fundamentally broad and unconventional in nature, it focuses on the creation and revision of protocols for storing and analyzing large quantities of data safely and efficiently. While most scientific fields such as biology, economics, mathematics, and physics produce large amounts of data, computer science focuses on the storage, analysis, and retrieval of the data. Computer scientists have made great advancements in all of these fields, but as they collect more and more data, scientists are posed with the issue of having to compute far more efficiently. To take a look at how this field conducts, records, and presents its studies, this paper will examine research articles to show how the field currently conducts research, prepares it for publication, and disperses it amongst its scholars.

Conducting Research

The Computer Science and Engineering discipline tends to engage in research across all different types of fields and is highly effective in terms of multidisciplinary collaboration. This allows researchers to pinpoint a problem and develop an innovative solution. Research in this field generally range in the form of experimentation, simulation, data collection, and prototyping.

Computer technology can process more data, store more information, and make it available to all. This discipline is connected with other fields of research because computers form the core for information analysis in most fields. In any science discipline, computers help scientist process and store collected data on their research. In addition to storing information, computer also help scientist monitor the instruments needed for data collection. Computer technology is also helpful to engineering disciplines such as social sciences where computers helps individuals connect with the global world and access large amount of information. An individual can also use computers to access different platforms of media.

Perhaps the most well-known computer science paper, "MapReduce: Simplified Data Processing on Large Clusters", authored by Jeffrey Dean and Sanjay Ghemawat and published by Google, highlights a fundamental concern with an efficient way to process big data. MapReduce, a now a very important concept for algorithms, has been cited over sixteen thousand times [1]. While the paper utilizes standard data reporting techniques such as graphs and figures, the data in those graphs serve a much different purpose than simple statistical data. Whereas experimental data is the foundation for the most research, the data found in MapReduce and other computer science papers only serves as supporting evidence for a larger theory. The

research that Dean and Ghemawat conducted consisted of coming up with the MapReduce algorithm, implementing it, and conducting analysis and testing to prove its worth.

One example of a computer engineering research paper is the relatively unknown "State Considered Harmful, A Proposal for a Stateless Laptop" by Joanna Rutkowska of the Invisible Things Lab. Her paper stresses the security flaws of modern CPU architecture, and is markedly less mathematical and lacks any quantifiable data. Instead of graphs, the essay consists of diagrams of computer systems peculiarly drawn as ASCII art (American Standard Code for Information Interchange, art using numbers, characters, and symbols) [2]. Rutkowska runs through the abstracted technical specifications of a typical laptop and argues for her solution to the aforementioned security flaws: laptops that don't carry state, i.e. having all data stored on a secure external device, not on the laptop. This paper is a follow up to the author's previous paper dealing with revealing the security flaws which this paper addresses. As Rutkowska described herself, these papers were born after her reading of a technical book on features of the architecture. Research in this case consisted of learning the technical specifications of a CPU architecture, and synthesizing a solution to cover up its security flaws.

Apart from pure computational theory, computer science is fundamentally an interdisciplinary field of study, since it allows large quantities of data to be computed in a relatively short period of time. Because of its interdisciplinary nature, computer science research papers often include data, graphs, and visualizations to help scholars of other disciplines get a better glimpse of the technical CSE theory at work. A key example of its interdisciplinary foundation lies within the research paper "MATtrack: A MATLAB-Based Quantitative Image Analysis Platform for Investigating Real-Time Photo-Converted Fluorescent Signals in Live Cells" authored by Jane Courtney, Elena Woods , Dimitri Scholz , William W. Hall , and Virginie W. Gautier. These researchers were able to design a development tool/program to generate maps and visualize protein trafficking patterns and regions of accumulation, all of which require new algorithms and excessive computational power [3]. Now, it should be noted that while the paper highlights the application of MATtracks in biology, the article focuses on the efficiency of algorithms to make sense of the biological data. It is using this data to show the biology researcher what improvements MATtracks has made in the extraction of data from raw biological data in the form of images [3]. In places where the paper is explaining the algorithms efficiencies, the authors of the paper have placed diagrams to show the algorithms process in making sense of the images.

Information Dispersal

Research and publications in the field of computer science and engineering are dispersed through many venues including the internet, academic journals, and research conferences. While most fields of study place prestige on research publications found in journals because of their stringent peer review process, computer scientists have stayed accustomed to conference publications. Although there is less emphasis on an established peer review process, we find that

the same concept still applies as institutions and corporations will self-review any material they produce and distribute. In order to maintain their excellent/well respected standing, they must adhere to standard factchecking. While research conferences were essential to the dispersal of early research, many scientists are now pushing the move towards publications through research journals.

While most scientific fields favor journal publications because of their thorough peer-review process, Journal publications require peer reviewers who are "senior scientist" or respected individuals within their field which did not exist up until recently. This new trend stands as a large amount of research articles to date are written independently and posted on the internet on personal or company blogs, only to be later submitted to conferences and published. This process which can be nearly considered self-publishing because of their availability on the world-wide-web, has been very efficient for the dispersal of academic work. Just recently members of the scientific community began proposing the switch from conference publications to journal publications because of the greatly beneficial peer-review process. As Lance Fortnow, professor of electrical engineering and computer science at Northwestern University, states in his paper "Time for Computer Science to Grow Up", The field of computer science "has matured and the conference model has fractured the discipline and skewered it towards short-term deadline-driven research" [4]. He states here that the conference style publications are now doing more harm than good by forcing a deadline on researchers to complete a work, rather than having the piece of work iteratively reviewed by experts in the field.

This culture of getting work done for a deadline and not returning to refine the work is common within the field and emerges as early as a student's undergraduate years at events such as hackathons. Hackathons are conference-like competitions in which participants are given a set amount of time, commonly ranging from half a day to an entire weekend to produce a tangible product for review. This culture of deadline driven work highlights the preference for agile conference publications over stringent and lengthy journal publications.

Format

The Computer Science and Engineering format, in order to maximize simplicity and reduce clutter, adheres to a less strict style of citation. This style, COSE (Council of Science Editors) prioritizes easy fact checking as it uses superscripts to link to citations. There are two main approaches that are in use today, COSE name-year system and COSE citation-sequence system [5]. The most prevalent is the citation-sequence system as it allows the researcher to simply denote each source with a number and use that as a marker for each source. Also, the references list does not need to be in alphabetical order. This saves time and space as most will generally not need to remember each citation until they need to look more into that topic or research. The name-year system is also simple in the fact that it follows a parenthetical citation format similar to APA, (Last Name, year).

COSE papers are generally formatted in ways where it is the most consumable and timesaving for the reader. This means that research papers may also include special keywords underneath the title in order to emphasize the intended readers/content. From this point forward, the format begins to look like an engineering or scientific report. In order, the paper consists of abstract, introduction, methods, results, discussion and references. These sections are easy to locate as are bolded and possibly italicized in the paper so readers can skim or skip any unnecessary information

The abstract is a section for readers who are looking to see if the article is something that they want to research about. The introduction will explain not only what the subject is about but even define specific words and certain concepts that a person in the discipline might not understand because the subject requires that individual to concentrate solely on the concept. This section serves to highlight the core concepts and what should be received/understood from the paper. This is also an excellent way to avoid clutter and make the paper concise.

Since a majority of research in Computer Science is theoretical/idealistic, The Methods and Results sections serve more to explain the ideas and concepts behind a certain works rather than provide statistical data. The methods section, can also go into the "possible solutions" section for Computer Science and Engineering papers.Within these papers, the authors generally add more commentary about the topic and include personal opinions as well. The author might talk about how some people may find the methods and results as useless or outdated. This section is then where the author may then explain why a certain method is the best solution for a certain problem.

Finally, the references section is similar to the work-cited page of a paper in MLA. There are three systems for citing within the text, citation-sequence, name-year, and citation-name. There is no universal preference for which citation system to use, however, the most notable difference between these systems is how it appears in the end references page. However, a specific citation style that a majority of Computer Science and Engineers use is the Association of Computing Machinery (ACM) citation method. In this method, the end reference section uses the cited author's' last name in alphabetical order. The difference between this format and other academic formats is that the URL is significant in the citation. Most of the research and scholarly information are most likely found on the internet rather than a book.

References

1. Dean, Jeffrey, and Sanjay Ghemawat. "Mapreduce: Simplified Data Processing On Large Clusters." Communications Of The ACM 51.1 (2008): 107-113. Academic Search Complete. Web. 9 Feb. 2016.

2. Rutkowska, Joanna. State Considered Harmful. [Internet]. Warsaw, Poland. N.e. The Invisible Things. [2015 Dec 26., cited 2016 Feb 9]. Available from: http://blog.invisiblethings.org/2015/12/23/state_harmful.html

3. Courtney, Jane, et al. "Mattrack: A MATLAB-Based Quantitative Image Analysis Platform For Investigating Real-Time Photo-Converted Fluorescent Signals In Live Cells." Plos ONE 10.10 (2015): 1-14. Academic Search Complete. Web. 9 Feb. 2016.

4. Fortnow, Lance. "Viewpoint Time for Computer Science to Grow up." *Communications of the ACM Commun. ACM* 52.8 (2009): 33. Web.

5. Carey P. 2014. Citing Your Sources: CSE (Council of Science Editors) Style. [Internet]. N.e. Boston(MA):Boston University Libraries; [N.d., cited 2016 Feb 14] . Available from: http://www.bu.edu/library/guide/csestyle/