UNIVERSITY OF WATERLOO

Dr. Wojciech Golab, Associate Professor
University of Waterloo, Canada
wgolab@uwaterloo.ca

March 8, 2022

Reviewer's opinion
on Ph.D. dissertation authored by
Maciej Kokociński
entitled:
Correctness of Highly-Available Eventually-Consistent Replicated Systems

1. Problem and its impact

The fundamental scientific problem discussed in the dissertation is the problem of *understanding* the correctness properties of a very popular class of data storage systems. In this context, "understanding" can be interpreted as formalizing and comparing correctness properties, as well as deciding whether such properties can be realized under specific combinations of assumptions regarding the data storage system and its environment.

The contributions of the dissertation have definite potential for positive impact on the practice of distributed computing, and also on academic research. For example, the impossibility results presented in the dissertation can help software engineers understand the operational limits of data storage systems, and guide design decisions during system development. Similarly, the correctness properties analyzed in the dissertation can be cited in product manuals to describe the behaviors of a system under different configurations of tuning parameters.

In an academic context, the new formal models and correctness specifications introduced in the dissertation have the potential to inspire further academic research, for example in the form of novel distributed protocols, analyses of correctness, and empirical performance evaluations. Indeed, the conference and journal publications arising from the dissertation are already starting to garner citations according to Google Scholar, a popular bibliometric tool.

2. Contribution

The main contribution is a rigorous analysis of mixed consistency systems, which support both weakly consistent highly available operations and strongly consistent operations that lack high availability. Specifically, the dissertation presents a new abstract data type called the *Acute Cloud Type* (Chapter 2), a formal framework for analyzing mixed consistency systems including a novel correctness criterion called *fluctuating eventual consistency* (Chapter 3), an impossibility result that captures the key limitation of such systems (Chapter 4), an extended formal framework that explicitly accounts for failures (Chapter 5), formal definitions of so-called session guarantees (Chapter 6), as well as a rigorous analysis of storage systems under different failure models and definitions of novel failure-aware correctness criteria (Chapter 7).

A substantial portion of the dissertation has been published in competitive scientific venues including the ACM Symposium on Principles of Distributed Computing (PODC) and the IEEE Transactions on Parallel and Distributed Systems (TPDS). According to the Computing Research and Education Association of Australasia[1] (CORE) ranking system for conferences and journals, ACM PODC and IEEE TPDS are both exceptionally competitive venues (A* rank, top 5-10%).

3. Correctness

The dissertation introduces both novel mathematical definitions (e.g., of desirable correctness criteria or undesirable behavioral anomalies), and theorems that make precise statements regarding such definitions. The definitions themselves are presented without proofs, which is standard practice in computer science, and are supported instead with intuitive explanations and illustrative examples (e.g., Figure 2.1). I am satisfied on the basis of these explanations and examples that the definitions meaningfully capture the intended behaviors. The theorems, on the other hand, are established using rigorous proofs. The chosen proof techniques are appropriate given the context, mostly based on first order logic and indistinguishability arguments, both of which have been used extensively in distributed computing theory. The "proof by contradiction" style is applied effectively to establish impossibility results. The level of detail in the proofs meets or exceeds what would be expected of a doctoral student at the University of Waterloo.

[1] https://www.core.edu.au/

4. Knowledge of the candidate

Background information on the subject of distributed computing is presented at several points in the dissertation. Chapter 2 provides a technical overview of several background topics, including the widely-cited Bayou system, basic eventual consistency (BEC), various progress guarantees, broadcast primitives, and design principles for distributed storage systems. Chapter 3 further provides a review of fundamental mathematical abstractions needed for the analysis presented in the dissertation, and also defines common correctness criteria within this mathematical framework. There is also a thorough review of related work in Chapter 8 that positions the contributions within the broader perspective of distributed storage systems. The background material in these chapters covers selected areas of discrete mathematics, logic, and computer science. The content is generally well written and covers sufficient scope. The dissertation cites more than one hundred references, which is fairly typical for a doctoral dissertation in computer science, and provides appropriate coverage of both recent developments (last 10 years) and classic research works. Overall, the thesis demonstrates a deep theoretical understanding of Information and Communication Technology, focusing on distributed data storage systems.

5. Conclusion

Taking into account what I have presented above and the requirements imposed by Article 13 of *the Act of 14 March 2003 of the Polish Parliament on the Academic Degrees and the Academic Title* (with amendments)[2], my evaluation of the dissertation according to the three basic criteria is the following:

A. Does the dissertation present an original solution to a scientific problem? (the selected option is marked with **X**)

[X]	[ ]	[ ]	[ ]	[ ]
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

B. After reading the dissertation, would you agree that the candidate has general theoretical knowledge and understanding of the discipline of **Information and Communication Technology**, and particularly the area of ....?

[X]	[ ]	[ ]	[ ]	[ ]
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

C. Does the dissertation support the claim that the candidate is able to conduct scientific work?

[X]	[ ]	[ ]	[ ]	[ ]
Definitely YES	*Rather yes*	*Hard to say*	*Rather no*	*Definitely NO*

[2] http://www.nauka.gov.pl/g2/oryginal/2013_05/b26ba540a5785d48bee41aec63403b2c.pdf