Natalia Kushik
Services répartis, Architectures, Modélisation, Validation, Administration des Réseaux (SAMOVAR)
Télécom SudParis
Institut Polytechnique de Paris
19 Pl. Marguerite Perey, 91120 Palaiseau
Tel. : +33 1 75 31 44 19 / e-mail : natalia.kushik@telecom-sudparis.eu

PhD thesis report for the manuscript of

Jakub Ruszil

entitled

Synchronization of Finite Automata - Problems and Generalizations

Prepared at Jagiellonian University

The thesis of Jakub Ruszil is devoted to a well known problem in automata theory - automata synchronization. The relevant decision problem is stated as follows: given an automaton, one needs to check if there exists a synchronizing sequence (SS or a synchronizing word) for it. The derivation problem requests to produce such a word for a given automaton. This problem has a number of applications, in the area of Computer Science in various domains, involving setting a system to a known current / initial state, such as for example, logic synthesis, verification and testing, etc.

The manuscript "Synchronization of Finite Automata - Problems and Generalizations" is well structured and is written in English. It is composed of 7 chapters, 81 pages in total, including the references.

Chapter 1 contains the Introduction of the thesis. It starts with the motivating example, illustrating the properties of synchronization over directed graphs and its utility. A brief presentation of the current state of the art is given afterwards, referring to the important Černý Conjecture. Finally, the thesis objectives are stated, with the detailed explanations of the latter, including their practical applicability. The thesis structure, facilitating the further reading, is also presented in this chapter.

Chapter 2 introduces the reader to the problem of synchronization and contains various important blocks. First of all, it is a background section with the necessary notions and notations. Secondly, the Černý Conjecture is discussed here too, and a detailed state of the art for the DFAs, existence check and derivation of SSs for them, the length of related sequences, the classes for which the Conjecture is proven, etc. are presented furthermore. I particularly liked how the author summarizes the main existing results in the area, which is not an easy task seeing how the problem has remained open for more than half of a century. Finally, this chapter also introduces a number of application scenarios for the synchronization problem, i.e., where setting a system to a known initial state can bring certain benefits. Examples include coding theory, model based testing, robotics, etc.

Chapter 3 is devoted to studying the careful synchronization, i.e., SSs for partial automata or PDAs. The author starts by presenting a class of PDAs with the shortest synchronizing word of length of the order $3^{(n/3)}$, improving the known result of Martyugin through reducing the size of the input alphabet. This is a very valuable contribution of the thesis, the only comment that I can give on it is that maybe more intuition about the construction of the machine itself could be given; the latter could simplify the reading and understanding. Further, there comes a proposal of a generic case of a 3-letter automaton with the length of an SS which is a power of e. The latter follows the proposed construction of an automaton A_p, $p = p=(k_1, .., k_s)$, with a shortest SS of length $lcm(k_1, .., k_s) + 1$. Similar to the previous case, maybe more explanations about the derivation of this generic case of the machines could be given. Quick question to

the author - for these machines is there a single SS with which is shortest ? This question can be applied to other sections and chapters as well, as I have noticed the author mostly uses "the" article for the shortest synchronizing word, does this mean that in all cases there is a single one ? (Otherwise, "a" shortest could be applicable).

Chapter 4 is devoted to studying another specific class of automata, namely automata with coinciding cycles. For this class of automata, it is shown that the length of a synchronizing sequence has a quadratic upper bound, which is of course, a very valuable contribution, that aligns with the studies in the area of Černý Conjecture. Interestingly, these automata are always synchronizable which is also proven in this chapter. Note that the chapter itself starts with the motivation of the usage of this class - I appreciate this part, however I am not totally convinced. Maybe, real-life specifications allowing such kinds of automata constructions could serve as one the directions for future work. At the same time, as the new knowledge of a new class with the quadratic upper bound on the SS length, this class is totally interesting.

Studies on particular automata classes continue in Chapter 5. Namely, the author moves from DFA to PFA, and checks the PFAs with the exponential upper bounds on their synchronizing words. One-cluster automata fall into this category and are the ones being investigated. First, the exponential upper bound on the length of the corresponding SS is proven, and later on the NP-hardness of the corresponding existence check of an SS for the one-cluster automata, is proven. The latter is made through a 3-SAT reduction, i.e., through constructing (in polynomial time) an automaton A_Φ from a given formula Φ; the former is synchronizable if and only if the latter is satisfiable. Note that the exponential upper bound is proven for an automaton with the exponential size of the input alphabet. May we assume, taking into account the second result, that the upper bound on the length of an SS for one-cluster automata can be smaller then ?

Chapter 6 describes an application scenario related to the complexity of careful synchronization. In particular, the author proposes a public key cryptosystem based on a number of PFAs, all having the same synchronizing word. The encryption takes a binary word as a plaintext and transforms it into an automaton representing a ciphertext. That is a very interesting attempt, given the complexity of finding such a word, indeed. However, there are a number of problems generating the automata in question. An option could be to rely on the results of the previous chapter and to go through the one-cluster automata for which a 3-SAT reduction, proving the NP-hardness of the related problem, has been already established. What I did not get well from the presentation of the cryptosystem is the set of keys and the procedures for each of the interested parties: do Alice and Bob share the same private / public key or are they different ? Alice is supposed not to know the private one of Bob, but if she knows how to synchronize the automata in question then she probably does ? This application scenario seems to be promising and needs to be further investigated, also from the practical point of view - in fact, I do not know how efficient it is to represent and send a ciphertext in the channel in the form of an automaton ? I encourage the author to continue his research on this subject, of course.

Chapter 7 concludes the work, summarizing the main contributions of the thesis. It would probably be interesting to include and discuss the future work in this part too, however I admit that these issues were listed in the chapters and sections directly.

The thesis itself presents a purely fundamental work and extends the automata theory, notably through studying particular automata classes with the tight upper bounds on the length of synchronizing words. The results of this work were published in the top ranked TCS conferences, such as CIAA and DLT, which are great achievements of the author. It could probably be beneficial also to note in the relevant chapters which contributions were published where (for example, in the Summary after the sections that the author provides).

Overall, the manuscript is nice to read and it showcases and combines various results of the author in the area of DFA and PFA synchronization, together with the particular applications of the problem and relevant length and complexity estimation. What is slightly missing in the presentation and organization, in my opinion, is the connection between the chapters to make it look like a "bigger picture" rather than a collection of the obtained results. Some motivation related questions could also serve as preambles - why this class is introduced, why is it studied, and how it brings us to another chapter where other types of automata are considered ?

Note however, that the questions and comments given above, do not affect the contributions of the author, on the contrary, some of them probably open new avenues for future work.

For all these reasons, I am convinced that the author, **Jakub Ruszil**, deserves the title of the **Doctor of Philosophy**, of Jagiellonian University.

07/01/2025

Natalia KUSHIK

Institut Polytechnique de Paris
Télécom SudParis
Institut Mines-Télécom
19 rue Marguerite Perey
91120 Palaiseau
France
www.telecom-sudparis.eu