MASARYKOVA UNIVERZITA

Annex No. 11 to the MU Directive on Habilitation Procedures and Professor Appointment Procedures

Habilitation Thesis Reviewer's Report

Masaryk University	
Faculty	Faculty of Informatics
Procedure field	Informatics
Applicant	**RNDr. Tomáš Masopust, Ph.D., DSc.**
Applicant's home unit, institution	Palacky University in Olomouc
Habilitation thesis	Partially Ordered Automata - Expressivity, Complexity, and Applications
Reviewer	**Prof. Dr. Henning Fernau**
Reviewer's home unit, institution	Theoretische Informatik, Universität Trier, Germany

Summary

T. Masopust has handed in a quite impressive survey of his more recent research, concerning many complexity aspects of (sub-)regular languages. Parts of his work has been presented at highly esteemed conferences like ICALP and MFCS. The research is of theoretical nature, but has quite a number of nice practical motivations, including concrete implications for practitioners. For instance, if (starting with the first pages) certain operations on languages (even from sub-regular language classes) can lead to (provably) very large automata, then these cannot be handled by standard applications that are based on these automata. In order to avoid the use of such state explosion, the lesson is to avoid the mentioned operation, which might lead (for instance) to a re-design of a domain-specific programming language.
Also, the general line of research to consider sub-regular language classes (that again are often well motivated from practical considerations), looks like a promising venue, also for future research. It parallels a bit the attempt in graph algorithms to study specific graph classes when facing NP-hard algorithmic questions. Unfortunately, as T. Masopust shows in several papers, the question itself if a certain regular language belongs to a certain sub-regular language class often turns out to be a hard question, unlike most of the prominent graph classes (to continue with this analogy). This might motivate a question to look for practically relevant sub-regular graph classes where membership in this class can be determined efficiently, possibly maintaining further positive algorithmic properties (like efficient decidability of equivalence or universality). In the habilitation thesis, the focus is rather set on well-known automata classes, for instance members of the Straubing-Thérien hierarchy.
Finally, several applications are presented, from database theory (more specifically, from the theory of semi-structured documents), from system theory (more precisely, concerning the notion of detectability in discrete event systems, or DES for short) and from privacy / security analysis (of DES).

Reviewer's questions for the habilitation thesis defence

Regarding separability, you considered the subsequence order (which is a wqo) and the suffix and prefix orders (which are not wqo),. The complexity results differ considerably. What about the infix (subword, or factor) ordering (also not a wqo), or possibly even lexicographic or length-lexicographic orderings? You also looked into "towers of subsequences" and "tower of prefixes". As "tower of infixes" would somehow interpolate between the mentioned two, it might be interesting to look into this matter. What do you know about it?

Concerning the complexity results that you obtained for subsequence and suffix orders: Only in two cases, you claim complexity results that are "complete"; is it true that the exact complexity status (completeness for a certain class) is still open? What are your conjectures in these cases? For instance, could it be that some of the PTIME-results are rather NL-completeness results?

In your discussion of Simon's equivalence $\sim_k$, it would be interesting to know how fast one could determine a maximum k such that u $\sim_k$ v holds for given u,v. What do you know about this problem?

What about decidability questions like: "Given two automata (DFA or NFA) A and A' and an integer k>0, does there exist a sequence of words $w_1, \ldots, w_k$ forming an R-tower between L(A) and L(A')?" Here, R should denote an ordering, like prefix, subsequence etc.

On page 40, you write that NL is the class of problems that can be efficiently parallelized. The more traditional opinion seems to be, however, that NC^1 is the class of problems that is efficiently parallelizable. In a sense, if you believe in NC^1\neq NL, one could even say that NL-complete problems (as the one from Theorem 82) is not efficiently parallelizable. Maybe, you can say more about your opinion here.

In all the problems you looked into, have you thought about special forms of poDFAs or poNFAs like those having parallel-sequential graphs with loops as underlying structure?

A more specific question: In the proof of Thm. 20 of your TCS paper with M. Thomazo from 2017, you refer to a paper from Hunt III and Rosenkrantz for certain properties of bitstrings that are "as hard to decide as universality" according to that paper. Could you explain in more detail how you arrive at this nice result from what Hunt III and Rosenkrantz write? At least, that paper contains no theorem that reads as you write it.

Conclusion

The habilitation thesis entitled "Partially Ordered Automata - Expressivity, Complexity, and Applications" by Tomáš Masopust ***fulfils*** requirements expected of a habilitation thesis in the field of Informatics.

In Fell, on March 3rd , 2021