THEORIE-
TAG 2023
A. Lin, G. Zetzsche (Eds.): Theorietag 2023, Kaiserslautern, 4. – 6.10.2023
RPTU & MPI-SWS Kaiserslautern, Technical Report, pp. 16–18.

Synchronization and Diversity of Solutions

Emmanuel Arrighi[(A)] Henning Fernau[(B)]
Mateus de Oliveira Oliveira[(C)] Petra Wolf[(D)]

[(A)]Universities of Bergen, Trier, Lyon
emmanuel.arrighi@gmail.com

[(B)]University of Trier
fernau@uni-trier.de

[(C)]Universities of Bergen, Stockholm
Mateus.Oliveira@uib.no

[(D)]Universities of Trier, Bergen, Bordeaux
mail@wolfp.net

This is a very short account on our paper presented this year at AAAI, see [2]. This paper considers the notion of diversity of solutions in the context of synchronizing words.

1. Introduction

A word w is said to be *synchronizing* for a DFA A if there is some state q of A such that any state q' is sent to q by w. The most elementary problem is to determine whether a given DFA has a synchronizing word. This can be decided in polynomial time. Nevertheless, in several applications, one is interested in finding a synchronizing word satisfying certain additional constraints. Here, the complexity landscape changes drastically: even determining the existence of a synchronizing word satisfying additional length or regularity constraints is NP-hard, see [6, 7] for some examples.

As not all important features of a solution may be formalized completely, several applications then request the enumeration of all solutions, possibly additionally satisfying for instance some form of minimality. In the context of strings (as solutions), several possibilities exists to define minimality, based on different partial orderings, like prefix, infix, or subsequence orders. In relation with synchronizing words, these have been discussed in [8]. There, it was shown that, for instance concerning the subsequence ordering $|$, the question if, given a DFA A and a word $u \in \Sigma^*$, there exists a $|$-minimal synchronizing word w with $u|w$ is NP-hard. This rules out at least a simple way to obtain enumeration algorithms that achieve polynomial delay, which implies that a data analyst might have to wait exponentially long even between seeing two different solutions enumerated. This is clearly not acceptable.

Support by the Research Council of Norway, DAAD and DFG is gratefully acknowledged.

Alternative notions have been proposed to overcome this problem, one of them being *diversity*, suggested in [3]. The idea is to find a *small* set of solutions that are sufficiently diverse from one another. How can we adapt the framework of solution diversity to the context of synchronization? One problem is that usual notions of diversity of solutions based on Hamming distance are not appropriate to measure diversity between strings. For instance, distinct solutions may have distinct length. Even strings of the same length that are very similar to each other on an intuitive level may have very large Hamming distance, as $w = abab...ab$ and $w' = baba...ba$ show. Therefore, we base our diversity measure on the notion of *edit distance*. Unfortunately, a set of solutions S in which any two of them are far apart from each other may still not capture solution diversity in our context, as if w is a synchronizing word, then any xwy is synchronizing. We therefore require that each word in S is subsequence-minimal synchronizing.

2. Our Results

The subsequence minimality requirement combined with edit distance not only guarantees that solutions in any given subset are genuinely distinct, but also provides a way of tackling diverse synchronization problems using the machinery of finite automata theory. On the one hand, Higman's lemma [9] implies that the set of subsequence-minimal synchronizing words in the language of an automaton is always finite. On the other hand, the computation of the edit distance between two words is a process that can be simulated using finite automata. More specifically, it is possible to construct finite automata accepting a suitable encoding of pairs of words that are far apart from each other.

Note that subsequence-minimal synchronization problems involving a single DFA A are already very hard. First, subsequence-minimal synchronizing words for a DFA A may have exponential length on the number of states of A. Second, determining if a given word w is subsequence-minimal among all synchronizing words in the language of a DFA A is co-NP-hard. Third, determining if a DFA A has two distinct subsequence-minimal synchronizing words is NP-hard. Finally, the problem of counting the set of subsequence-minimal synchronizing words is #P-hard. We also remind the reader of the already mentioned NP-hardness result concerning the extension problem variant [8].

In order to cope with the inherent intractability of synchronization problems, we leverage on the framework of parameterized complexity theory [5]. In particular, we show that for each fixed value of r, interesting computational problems requiring a diverse set with r subsequence-minimal synchronizing words can be solved in time that is fixed parameter tractable with respect to the size of the synchronizing automaton A. Previously, algorithms with an FPTdependence in $|A|$ were unknown even for $r = 2$. Using our approach, we also show that given a DFA A with state set Q over an alphabet Σ, and a word $w \in \Sigma^*$, one can determine in time $O(f(|\Sigma|, |Q|) \cdot |w|)$, for some function f, if some subsequence-minimal synchronizing word for A is a subsequence of w, and we can construct such a subsequence in case the answer is affirmative. Recall that the unparameterized version of this problems is already co-NP-hard. Our main algorithmic result states that, given numbers $r, k \in \mathbb{N}$, a DFA A, and a possibly nondeterministic finite automaton B over an alphabet Σ, the problem of computing a subset $\{w_1, \ldots, w_r\} \subseteq L(B)$ of subsequence-minimal synchronizing words for A, with pairwise edit distance of at least k, can be solved in time $O(f_A(r,k) \cdot |B|^r \log(|B|))$ for some suitable function f depending only on A, r and k. Intuitively, the automaton A is a specification of a system which we want to synchronize (or reset), and B is a specification of the set of words that are allowed to be used as synchronizing sequences. As stated above, the unparameterized version of this problem is NP-hard even if we are interested in finding a single solution and the language of the automaton B is as simple as ab^*a. As a consequence of our main result, given a word $w \in \Sigma^*$, the problem of determining whether there exist r subsequence-minimal synchronizing words for A that are subsequences of w and that are at least k apart from each other can be solved in time $O(f_A(r,k) \cdot |w|^r \log(|w|))$.

It turns out that our notion of diversity of solutions can be applied in other contexts where solutions are strings whose sizes may have vary. We adapt our framework to the realm of conformant planning, where the goal is to design plans that achieve goals irrespectively of initial conditions and of nondeterminism that may occur during the execution of these plans [1, 4, 10]. Throughout our paper, classical automata constructions help prove our results.

References

[1] A. S. ANDERS, *Reliably arranging objects: a conformant planning approach to robot manipulation*. Ph.D. thesis, Massachusetts Institute of Technology, USA, 2019.

[2] E. ARRIGHI, H. FERNAU, M. DE OLIVEIRA OLIVEIRA, P. WOLF, Synchronization and Diversity of Solutions. In: *The Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI-23)*. AAAI, 2023, 11516–11524.

[3] J. BASTE, M. R. FELLOWS, L. JAFFKE, T. MASARÍK, M. DE OLIVEIRA OLIVEIRA, G. PHILIP, F. A. ROSAMOND, Diversity of solutions: An exploration through the lens of fixed-parameter tractability theory. *Artificial Intelligence* **303** (2022), 103644.

[4] B. BONET, Conformant plans and beyond: Principles and complexity. *Artificial Intelligence* **174** (2010) 3-4, 245–269.

[5] R. G. DOWNEY, M. R. FELLOWS, *Parameterized Complexity*. Springer, 1999.

[6] D. EPPSTEIN, Reset sequences for monotonic automata. *SIAM Journal on Computing* **19** (1990) 3, 500–510.

[7] H. FERNAU, V. V. GUSEV, S. HOFFMANN, M. HOLZER, M. V. VOLKOV, P. WOLF, Computational Complexity of Synchronization under Regular Constraints. In: P. ROSSMANITH, P. HEGGERNES, J.-P. KATOEN (eds.), *44th International Symposium on Mathematical Foundations of Computer Science (MFCS 2019)*. Leibniz International Proceedings in Informatics (LIPIcs) 138, Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2019, 63:1–63:14.

[8] H. FERNAU, S. HOFFMANN, Extensions to minimal synchronizing words. *Journal of Automata, Languages and Combinatorics* **24** (2019), 287–307.

[9] G. HIGMAN, Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society (3)* **2** (1952) 7, 326–336.

[10] H. PALACIOS, H. GEFFNER, Compiling Uncertainty Away in Conformant Planning Problems with Bounded Width. *Journal of Artificial Intelligence Research* **35** (2009), 623–675.