Report on the thesis (disertační práce) of **Ondřej Klíma** submitted to the Masaryk University in Brno (CZ) to obtain a PhD

Miki Hermann
Chargé de recherche CNRS

Ondřej Klíma considers in his thesis the algorithmic problems of unification modulo the axioms of associativity and idempotence. More precisely, he considers several variants of the decision problem asking whether there exists a substitution σ that makes a finite system of term couples, written as a set of equations $S = \{s_1 =^? t_1, \ldots, s_n =^? t_n\}$, equal in the equational theory generated by the axioms E of associativity $(x \cdot y) \cdot z = x \cdot (y \cdot z)$ and idempotence $x \cdot x = x$. These problems are important for an arbitrary set of equational axioms E in automated deduction, in computer aided theorem proving, and they found recently also a significant impact in the domain of computer aided verification. Some 20 years ago, researchers in automated deduction started to develop unification algorithms for the most usual equational theories, like (combinations of) associativity, commutativity, idempotence, unit element, distributivity, etc. This quest for efficient unification algorithms immediately triggered the question for the computational complexity of the unifiability decision problem. The complexity of the unifiability problem for some sets of axioms, like associativity and commutativity, were solved quickly, whereas others resisted a long time, or their exact complexity could not be pinned down until now. One of these difficult axiom sets is also associativity and idempotence. Hence, the complexity questions considered by Ondřej Klíma in his thesis constitute a collection of open problems since at least 1985.

From an algebraic point of view, the unification problem modulo associativity and idempotence asks for satisfiability of equations in the variety of idempotent semigroups, also called *bands*. Ondřej Klíma determines in his thesis the computational complexity of several variants of this unifiability question. Chapter 1 of his thesis introduces the necessary background from algebra, complexity, and automated deduction, as well as the overall used notation. Chapter 2 considers the problem of AI-*matching*, i.e., the restriction of the unification problem to systems S where the right-hand side terms t_i are ground (without variables), in presence of only the associativity (A) and idempotence (I) axioms, where the signature consists only of the binary concatenation symbol $\cdot$ and constants (also called *elementary* AI-matching). Ondřej Klíma carefully proves first that this problem is NP-hard, by a polynomial reduction from the well-known decision problem 3SAT. A similar result for *general* AI-matching, i.e., in presence of additional uninterpreted functional symbols, was already known. However, Klíma's proof constitutes a stronger result, since it is done for the problem of *elementary* AI-matching. Although the proof could be considered as routine, the author performs it with great finesse. This proof carries over to the elementary AIU-matching problem, in presence of the additional unit element. In fact, the proof is done for the latter, followed by a remark that the presence of the unit element can be avoided, to get the proof of the former problem. However, the most important part of Chapter 2 is the NP-membership proof for the elementary AI-matching problem. Ondřej Klíma used the confluent and terminating rewrite system for idempotent semigroups, presented by Siekmann and Szabó in [5], to determine a polynomial upper bound on the size of the solutions, what immediately implies the NP-membership result and delivers therefore the NP-completeness result. This NP-membership result is of considerable difficulty and the author applied highly nontrivial methods to obtain the desired result. The results of Chapter 2 have been published in the paper [3].

Chapter 3 considers the AI-unification and AIU-unification problems, i.e., the unification problem modulo the axioms of associativity and idempotence, with the presence of the unit element in the latter case, respectively. The author treats both the elementary as well as the general AI- and AIU-unification problems. NP-hardness carries over from the matching results in Chapter 2. The difficulty is once more the NP-membership of the considered unification problems. Ondřej Klíma defines first the *simple* AIU-unification problem, based on the concept of a system with complete information about invariants. The author then shows that simple AIU-unification problems can be solved in polynomial time. The notion of simple AIU-unification problems, the development of the corresponding algorithm, and its analysis require highly nontrivial amount of invention and work. Based on the polynomial-time decidability of simple AIU-unification problems, the author derives immediately the NP-membership of the related AI- and AIU-unification problems. This is done by the usual technique of guessing, from a finite set of possibilities, a mapping that determines the complete information about invariants, followed by the aforementioned algorithm that verifies in polynomial time whether the guessed simple AIU-unification problem has a solution. The result on simple AIU-unification problems has also an impact on the AI-matching problems considered in the previous chapter. First, it implies a better upper bound on the size of solutions than the previous proof in Section 2.3. Second and more important, it proves the NP-membership also for the general AI-matching problem, what was missing in Chapter 2. The results of Chapter 3 have been published in the paper [2].

Chapter 4 considers the unification problem in proper varieties of bands, i.e., modulo a binary associative and idempotent function symbol that satisfies some additional equational axioms. Ondřej Klíma considers the lattice of varieties of bands (see pages 8 and 58) usually studied by algebraists. He determines first the complexity of elementary matching and unification problems in these varieties. It is interesting to see that a clear frontier between the tractable and intractable cases could have been determined for both matching and unification. Moreover, this borderline is dichotomic in the sense that the corresponding matching and unification problems are either polynomial or NP-complete. Except for one previously known result by Kapur and Narendran [1] concerning the elementary unification problem in commutative idempotent semigroups, all the presented results are new. The author derives afterwards similar dichotomic results for general matching and unification problems. Finally, he proves a generic theorem that allows to determine the complexity of the general matching and unification problems in the varieties of bands from the structure of the band's signature. The results of Chapter 4 have been accepted for publication in the International Journal of Algebra and Computation.

Chapter 5 considers the problem of solving systems of equations over a given fixed finite semigroup or band. Ondřej Klíma combines in this chapter previously known results, mostly those of [4], with several new results. The main result of this chapter proves that the satisfiability problem of a system S of equations over a given fixed band B is polynomial if B is a regular band, provided that the right-hand sides of the equations in S are ground terms, or if B is commutative; otherwise the satisfiability problem is NP-complete. The results of Chapter 5 are based on common work with Pascal Tesson and Denis Thérien. Since the work was done in a cooperative manner, some results from this chapter also naturally appear in the PhD thesis of Pascal Tesson.

The thesis is completed by a very brief conclusion, which is only one and a half pages long, therefore leaving the reader unsatisfied. I would have preferred a longer conclusion containing a comparison and a discussion, that would place the obtained results of Ondřej Klíma in a broader scope of automated deduction and verification, underlining this way the really important scientific impact of the thesis.

Ondřej Klíma presents an excellent thesis in which he solves several longtime open problems, for which many internationally renowned researchers in theoretical computer science were unable to find a solution. The presented results are deep and require perfect knowledge of the underlying algebra and complexity theory methods. Ondřej Klíma shows in his thesis that he is capable to perform a very deep and inventive research, developing highly nontrivial constructions, yielding excellent and detailed proofs. Klíma's presentation is very clear and pedagogic. The thesis constitutes a very important and complete document in the field of complexity in automated deduction. The author also shows that he is capable to work on his own, as well as in a fruitful collaboration with other colleagues.

For all the aforementioned reasons underlining the scientific qualities of Ondřej Klíma I strongly recommend to award him the title of PhD.

References

[1] D. Kapur and P. Narendran. Complexity of unification problems with associative-commutative operators. *Journal of Automated Reasoning*, 9(2):261–288, 1992.

[2] O. Klíma. Unification modulo associativity and idempotency is NP-complete. In K. Diks and W. Rytter, editors, *Proceedings 27th International Symposium on Mathematical Foundations of Computer Science (MFCS 2002), Warsaw (Poland)*, volume 2420 of *Lecture Notes in Computer Science*, pages 423–432. Springer-Verlag, August 2002.

[3] O. Klíma and J. Srba. Matching modulo associativity and idempotency is NP-complete. In M. Nielsen and B. Rovan, editors, *Proceedings 25th International Symposium on Mathematical Foundations of Computer Science (MFCS 2000), Bratislava (Slovakia)*, volume 1893 of *Lecture Notes in Computer Science*, pages 456–466. Springer-Verlag, August 2000.

[4] C. Moore, P. Tesson, and D. Thérien. Satisfiability of systems of equations over finite monoids. In J. Sgall, A. Pultr, and P. Kolman, editors, *Proceedings 26th International Symposium on Mathematical Foundations of Computer Science (MFCS 2001), Mariánské Lázně (Czech Republic)*, volume 2136 of *Lecture Notes in Computer Science*, pages 537–547. Springer-Verlag, August 2001.

[5] J. Siekmann and P. Szabó. A Noetherian and confluent rewrite system for idempotent semigroups. *Semigroup Forum*, 25:83–110, 1982.

Written in 2 copies:
one for the candidate,
one for the office of the dean.

In Paris, January 20, 2004.
Miki Hermann