Information and Computation ••• (••••) •••–•••

Contents lists available at ScienceDirect

Information and Computation

www.elsevier.com/locate/yinco

Finding all solutions of equations in free groups and monoids with involution[☆]

Volker Diekert[a], Artur Jeż[b,*,1], Wojciech Plandowski[c,2]

[a] Institut für Formale Methoden der Informatik, University of Stuttgart, Germany
[b] Institute of Computer Science, University of Wroclaw, Poland
[c] Institute of Informatics, University of Warsaw, Poland

ARTICLE INFO

Article history:
Received 3 December 2014
Received in revised form 21 August 2015
Available online xxxx

Keywords:
Word equations
Equations in free groups
String unification
Regular constraints
Equations in free monoids with involution

ABSTRACT

This paper presents a PSPACE algorithm which yields a finite graph of exponential size that describes the set of all solutions of equations in free groups as well as the set of all solutions of equations with rational constraints in free monoids. This became possible due to the recent *recompression* technique.

While this technique was successfully applied for pure word equations without involution or rational constraints it could not be used as a black box for free groups. Actually, the presence of an involution and rational constraints complicates the situation and some additional analysis is necessary. Still, the technique is general enough to accommodate both extensions. In the end, it simplifies proofs that satisfiability of word equations is in PSPACE and the corresponding result for equations in free groups with rational constraints. As a byproduct we can decide in PSPACE whether the solution set is finite.

1. Introduction

A word equation is a simple object. It consists of a pair (U, V) of words over constants and variables and a solution is a substitution of the variables by words in constants such that U and V become identical words. The study of word equations has a long tradition. Let *WordEquation* be the problem of deciding whether a given word equation has a solution. It is fairly easy to see that WordEquation reduces to Hilbert's 10th Problem (in Hilbert's famous list presented in 1900 for his address at the International Congress of Mathematicians). Hence in the mid 1960s the Russian school of mathematics outlined the roadmap to prove undecidability of Hilbert 10th Problem via undecidability of WordEquation. The program failed in the sense that Matiyasevich proved Hilbert's 10th Problem to be undecidable in 1970, but by a completely different method, which employed number theory. The missing piece in the proof of the undecidability of Hilbert's 10th Problem was based on methods due to Robinson, Davis, and Putnam [1]. On the other hand, in 1977 Makanin showed in a seminal paper [2] that WordEquation is decidable! The program went a different way, but its outcome were two major achievements in mathematics. Makanin's algorithm became famous since it settled a long standing problem and also because his algorithm had an extremely complex termination proof. In fact, his paper showed that the existential theory of equations in free

[☆] A preliminary version of this paper was presented as an invited talk at CSR 2014 in Moscow, June 7–11, 2014.
[*] Corresponding author.
E-mail address: aje@cs.uni.wroc.pl (A. Jeż).
[1] This work was supported under National Science Centre, Poland project number 2014/15/B/ST6/00615.
[2] This work was supported by the National Science Centre, Poland, grant no 2014/13/B/ST6/00770.

http://dx.doi.org/10.1016/j.ic.2016.09.009
0890-5401/© 2016 Elsevier Inc. All rights reserved.

monoids is decidable. This is close to the borderline of decidability as already the $\forall\exists^3$ positive theory of free monoids is undecidable [3]. Furthermore Makanin extended his results to free groups and showed that the existential and positive theories in free groups are decidable [4,5]. Later Razborov was able in [6] (partly shown also in [7]) to describe the set of all solutions for systems of equations in free groups (see also [8] for a description of Razborov's work). This line of decidability results culminated in the proof of Tarski's conjectures by Kharlampovich and Myasnikov. In a series of papers ending in [9] they showed: 1.) The elementary theory of free groups is decidable. 2.) Free non-abelian groups are elementary equivalent. The second result has also been shown by Sela [10], independently.

Another branch of research was to extend Makanin's result to more general algebraic structures including free partially commutative monoids [11,12], free partially commutative monoids with involution, graph groups (also known as right-angled Artin groups) [13], graph products [14], and hyperbolic groups [15,16]. In all these cases the existential theory of equations is decidable. Proofs used the notion of *equation with rational constraints*, which was first developed in the habilitation of Schulz [17]. The concept of equation with rational constraints is used also throughout the present paper.

In parallel to these developments there were drastic improvements in the complexity of deciding WordEquation. It is fairly easy to see that the problem is NP-hard. Thus, NP is a lower bound. First estimations for the time complexity on Makanin's algorithm for free monoids led to a tower of several exponentials, but it was lowered over time to EXPSPACE in [18]. On the other hand it was shown in [19] that Makanin's algorithm for the satisfiability of equations in free groups is not primitive recursive. (Already in the mid 1990 this statement was somehow puzzling and counter-intuitive, as it suggested a strange crossing of complexities: The existential theory in free monoids seemed to be easier than the one in free groups, whereas it was already known at that time that the positive theory in free monoids is undecidable, but decidable in free groups.) The next important step was done by Plandowski and Rytter, whose approach [20] was the first essentially different than Makanin's original solution. The main idea was to apply compression to WordEquation and the result was that the length-minimal solution of a word equation compresses well, in the sense that Lempel–Ziv encoding, which is a popular practical standard of compression, of such a solution is exponentially smaller than the solution itself (if the solution is at least exponential in the length of the equation). This yielded an $\mathsf{npoly}(n, \log N)$ algorithm for length n WordEquation with a length-minimal solution of length N, note that at that time the only available bound on N was the triply exponential bound by Makanin. Still, this result prompted Plandowski and Rytter to formulate a (still open) conjecture that WordEquation is NP-complete.

Soon after a doubly exponential bound on N was shown by Plandowski [21], this bound in particular used the idea of representing the solutions in a compressed form (in fact, the equation as well is kept in a compressed form) as well as employing a novel type of factorisations. Exploiting better the interplay between factorisations and compression Plandowski showed that WordEquation is in PSPACE, i.e., it can be solved in polynomial space and exponential time [22]. His method was quite different from Makanin's approach and more symmetric. Furthermore, it could be also used to generate all solutions of a given word equation [23], however, this required nontrivial extensions of the original method.

Using Plandowski's method Gutiérrez showed that satisfiability of equations in free groups is in PSPACE [24], which led Diekert, Hagenah and Gutiérrez to the result that the existential theory of equations with rational constraints in free groups is PSPACE-complete [25]. Without constraints PSPACE is still the best upper bound, although the existential theories for equations in free monoids (with involution) and free groups are believed to be NP-complete. Since this proof generalized Plandowski's satisfiability result [22], it is tempting to also extend the generator of all solutions [23]. Indeed, Plandowski claimed that his method applies also to free groups with rational constraints, but he found a gap in his generalization [26].

However in 2013 another substantial progress in solving word equations was done due to a powerful recompression technique by Jeż [27]. His new proof that WordEquation is in PSPACE simplified the existing proofs drastically. In particular, this approach could be used to describe the set of all solutions rather easily, so the previous construction of Plandowski [23] was simplified as well.

What was missing however was the extension to include free monoids with involution and therefore free groups and another missing block was the presence of rational constraints. Both extensions are the subject of the present paper.

Outline

We first follow the approach of [25] how to (bijectively) transform (in polynomial time) the set of all solutions of an equation with rational constraints over a free group into a set of all solutions of an equation with regular constraints over a free monoid with involution, see Section 2.2. Starting at that point in Section 3 we formulate the main technical claim of the paper: (effective) existence of a procedure that transforms equations over the free monoid and (roughly speaking) keeps the set of solutions as well as does not increase the size of the word equation; in particular in this section we make all the intuitive statements precise. Moreover, we show how this procedure can be used to create a PSPACE-transducer which produces a finite graph (of exponential size) describing all solutions and which is nonempty if and only if the equation has at least one solution. Moreover, the graph also encodes whether or not there are finitely many solutions, only. The technique of recompression simplifies thereby [25] and it yields the important new feature that we can describe all solutions. The next Section 4 is devoted to a proof of the statements from Section 3: we formalize, what type of factors we compress, see Section 4.2, and how to compress them. Lastly, in Section 5, we show that a slight modification of our method also provides an $\mathcal{O}(n^3 \log N)$ upper bound on the (non-deterministic) running time in case of a free group, where n is a size of the equation and N the length of the length-minimal solution.

Please cite this article in press as: V. Diekert et al., Finding all solutions of equations in free groups and monoids with involution, Inf. Comput. (2016), http://dx.doi.org/10.1016/j.ic.2016.09.009