The Complexity of Solving Word Equations

Dependable Systems - Dirk Nowotka

Project Description

A *word equation* is a formal equality $U = V$, where U and V are words (called the left and right side of the equation respectively) over an alphabet $A \cup X$; $A = \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \ldots\}$ is the alphabet of *constants* or *terminals* and $X = \{x_1, x_2, x_3, \ldots\}$ is the set of *variables*. A *solution* to the equation $U = V$ is a morphism $h : (A \cup X)^* \to A^*$ that acts as the identity on A and satisfies $h(U) = h(V)$; h is called the assignment to the variables of the equation. For instance, $U = x_1\mathtt{ab}x_2$ and $V = \mathtt{a}x_1x_2\mathtt{b}$ define the equation $x_1\mathtt{ab}x_2 = \mathtt{a}x_1x_2\mathtt{b}$, whose solutions are the morphisms h with $h(x_1) = \mathtt{a}^k$, for $k \geq 0$, and $h(x_2) = \mathtt{b}^\ell$, for $\ell \geq 0$. An equation is *satisfiable* (in A^*) if it admits a solution $h : (A \cup X)^* \to A^*$. A set (or system) of equations is satisfiable if there exists an assignment of the variables of the equations in this set that is a solution for all equations.

Makanin proved in 1977 that the satisfiability of word equations *is* decidable. More than a decade after Makanin showed that the satisfiability of word equations is decidable, the focus shifted towards identifying the complexity of solving word equations. Plandowski showed in 1999 that this problem is in PSPACE. Recently, Jez applied a new technique called recompression to word equations and obtained a a proof (in 2017) that the satisfiability of word equations can be decided in linear space. However, there is a mismatch between this upper bound and the known lower bound: we only know that solving word equations is NP-hard. It is a major open problem to show that solving word equations is in NP.

There are, however, several classes of word equations (defined by restrictions on their left and right side) whose satisfiability was shown to be in NP, or even in P. For instance, equations where the left and right side contain only repeated occurrences of a single variable and some terminals can be solved in polynomial time. Also, strictly regular ordered equations, where the sides contain exactly the same variables, in exactly the same order, but interleaved with different occurrences of strings of terminals, can be solved in NP-time.

The goal of this project is to identify new relevant classes of word equations and settle the complexity of solving them. For instance, we are interested in the complexity of solving equations from:

- the class of reversed ordered equations: the sides contain exactly the same variables, in reversed order;
- the class of quadratic equations, where each variable occurs in total, in both sides, at most twice;
- the class of equations where at most one variable occurs more than once in total, in both sides.

Applicable For

Bachelorstudents ☑
Masterstudents ☑

Keywords

Word Equations
Combinatorics on Words
Formal Languages
Complexity

Contact

Dr. Joel Day, Dr. Florin Manea
@ {jda,flm}@informatik.uni-kiel.de