The Satisfiability of Word Equations: Decidable and Undecidable Theories

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.

In order to solve Hilbert's tenth problem in the negative, Markov showed a reduction from word equations to Diophantine equations, in the hopes that word equations would prove to be undecidable. However, Makanin proved in 1977 that the satisfiability of word equations *is* in fact decidable. Though Markov's approach was unsuccessful, similar ones, based on extended theories of word equations, can also be explored. Matiyasevich showed in 1968 a reduction from the more powerful theory of word equations with linear length constraints (i.e., linear relations between word lengths) to Diophantine equations. Whether this theory is decidable remains a major open problem.

In recent years, deciding the satisfiability of systems of word equations has also become an important problem in fields such as formal verification and security where string solvers such as HAMPI, CVC4, Stranger, ABC, Norn, S3P and Z3str3 have become more popular. However, in practice more functionality than just word equations is required, so solvers often extend the theory of word equations with certain functions: the aforementioned linear arithmetic over the length, but also operators such as replace-all, extract, reverse, and various predicates such as numeric-string conversion predicate, regular-expression membership, etc.. Whether the theory of word equations enhanced with a length function is decidable is still a major open problem, but in the case of other operators (predicates) the decidability/undecidability of the satisfiability of word equations extended with the respective operators (predicates) was settled.

The goal of this project is to analyse the decidability of the satisfiability problem for word equations extended with new operators or predicates, inspired from practical applications.

Applicable For

Bachelorstudents ☑
Masterstudents ☑

Keywords

Word Equations
Combinatorics on Words
Formal Languages
Decidability

Contact

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