ELSEVIER

Theoretical Computer Science 297 (2003) 271–280

Theoretical Computer Science

www.elsevier.com/locate/tcs

Equations in free semigroups with involution and their relation to equations in free groups

Claudio Gutiérrez

Depto. de Ciencias de la Computación, Universidad de Chile, Blanco Encalada 2120, Santiago, Chile

Abstract

A free semigroup with involution (FSI) is essentially the set of words over a given alphabet plus an operator which reverses words. The paper introduces equations in FSI and show that they are the right objects to deal with when studying the complexity of equations in free groups. On these lines, we generalize to FSI several results valid for word equations, like the overlapping lemma, the $2^{\mathcal{O}(|E|)}$-bound on the exponent of periodicity of minimal solutions, and the NP-hard lower bound.

The main result of the paper is the reduction of the problem of satisfiability of equations in free groups to the satisfiability of equations in FSI by a non-deterministic polynomial time transformation.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: Word equation; Free group equation; Free semigroup equation

1. Introduction

The study of the problem of solving equations in free semigroups with involution (FSI), also called unification in FSI, and its computational complexity is a problem closely related to the problems of solving equations in free semigroups and in free groups, two problems which lately have attracted much attention of the theoretical computer science community [3,13,14,16,5].

FSI is a structure which lies in between that of free semigroups and free groups. Besides its rich relationship with semigroups and groups, the axioms defining a semigroup with involution show up in several important theories, like algebras of binary relations, transpose in matrices and inverse semigroups.

E-mail address: cgutierr@dcc.uchile.cl (C. Gutiérrez).

0304-3975/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0304-3975(02)00642-4

The problem of solving equations in free semigroups was proven to be decidable by Makanin in 1976 in a long paper [11]. Some years later, in 1982, again Makanin proved that solving equations in free groups was a decidable problem [12]. The technique used was similar to that of the first paper, although the details are much more involved. He reduced the problem of solving equations in free groups to finding special solutions (non-contractible) to equations in FSI, and showed that the latter problem is decidable. Both of Makanin's algorithms have received much attention. The enumeration of all unifiers was done by Jaffar for semigroups [7] and by Razborov for groups [15]. Then, the complexity became the main issue. Several authors have analyzed the complexity of Makanin's algorithm for semigroups [7,17,1], being EXPSPACE the best upper bound so far [3]. Very recently Plandowski, without using Makanin's algorithm, presented an upper bound of PSPACE for the problem of satisfiability of equations in free semigroups [14]. On the other hand, the analysis of the complexity of Makanin's algorithm for groups was done by Koscielski and Pacholski [9], who showed that it is not primitive recursive. With respect to lower bounds, the only known lower bound for both problems is NP-hard.

The paper introduces formally the concept of equation in FSI (a concept touched in Makanin's papers) and suggests it is a good mathematical object to study when dealing with the complexity of equations in free groups. In this framework, we prove that several results valid for word equations can be generalized to this new setting. The most important of these are Rytter–Plandowski's overlapping lemma (our Lemma 2), the fact that standard bounds on the exponent of periodicity of minimal solutions to word equations also hold with minor modifications in the case of FSI (Theorem 5), and the lower NP-hard lower bound on the complexity of solving equations in FSI.

On these lines, the main result of the paper is the reduction of the problem of solvability of equations in free groups to that of equations in FSI (Theorem 9). This goal is achieved by the generalizations mentioned above, using some of Makanin's results in [12], and showing how these results can be linked (Proposition 3).

The claims about the importance of this type of equations are confirmed by the developments after the submission of this paper. Using the results presented here, in [5] it was proved that solvability of equations in FSI is in PSPACE; hence, solvability of equations in free groups is in PSPACE. Afterwards this latter result was generalized in [2] to show that solvability of equations in free groups with rational constraints is PSPACE-complete.

For concepts of word combinatorics we will follow the notation in [10]. By ε we denote the empty word.

2. Equations in free semigroups with involution

A *semigroup with involution* (SI) is an algebra with a binary associative operation (written as concatenation) and a unary operation $()^{-1}$ with the equational axioms

$$(xy)z = x(yz), \quad (xy)^{-1} = y^{-1}x^{-1}, \quad x^{-1-1} = x. \tag{1}$$

A *free* semigroup with involution (FSI) is an initial algebra for this variety. It is not difficult to check that for a given alphabet C, the set of words over $C \cup C^{-1}$ together with the operator $()^{-1}$, which reverses a word and changes every letter to its twin (e.g. a to a^{-1} and conversely) is a free algebra for the SI over A.

Equations and solutions: Let C and V be two disjoint alphabets of constants and variables, respectively. Define $C^{-1} = \{c^{-1}: c \in C\}$. Similarly for V^{-1}. An *equation* E in FSI with constants C and variables V is a pair (w_1, w_2) of words over the alphabet $\mathscr{A} = C \cup C^{-1} \cup V \cup V^{-1}$. The number $|E| = |w_1| + |w_2|$ is the *length* of the equation E and $|E|_V$ will denote the number of occurrences of variables in E. These equations are also known as *equations in a paired alphabet*.

A map $S: V \to (C \cup C^{-1})^*$ can be uniquely extended to an SI-homomorphism $\bar{S}: \mathscr{A}^* \to (C \cup C^{-1})^*$ by defining $S(c) = c$ for $c \in C$ and $S(u^{-1}) = (S(u))^{-1}$ for $u \in C \cup V$. We will use the same symbol S for the map S and the SI-homomorphism $\bar{S}$. A *solution* S of the equation $E = (w_1, w_2)$ is (the unique SI-homomorphism defined by) a map $S: V \to (C \cup C^{-1})^*$ such that $S(w_1) = S(w_2)$. The length of the solution S is $|S(w_1)|$. By $S(E)$ we denote the word $S(w_1)$ (which is the same as $S(w_2)$). Each occurrence of a symbol $u \in \mathscr{A}$ in E with $S(u) \neq \varepsilon$ determines a unique factor in $S(E)$, say $S(E)[i, j]$, which we will denote by $S(u, i, j)$ and call simply an *image* of u in $S(E)$.

Equivalence relation (S, E): For a given word w, a *position* is an integer $p \in [1, |w|]$ (which indicates the position of a letter in the word). Let S be a solution of E and P the set of positions of $S(E)$. Define the binary relation $(S, E)'$ in $P \times P$ as follows: given positions $p, q \in P$, $p(S, E)'q$ if and only if one of the following hold:

1. $p = i + k$ and $q = i' + k$, where $S(x, i, j)$ and $S(x, i', j')$ are images of x in $S(E)$ and $0 \leqslant k < |S(x)|$.
2. $p = i + k$ and $q = j' - k$, where $S(x, i, j)$ and $S(x^{-1}, i', j')$ are images of x and x^{-1} in $S(E)$ and $0 \leqslant k < |S(x)|$.

Then define (S, E) as the transitive closure of $(S, E)'$. Observe that (S, E) is an equivalence relation.

Contractible words: A word $w \in \mathscr{A}^*$ is called *non-contractible* if for every $u \in \mathscr{A}$ the word w contains neither the factor uu^{-1} nor $u^{-1}u$. An equation (w_1, w_2) is called non-contractible if both w_1 and w_2 are non-contractible. A solution S to an equation E is called non-contractible if for every variable x which occurs in E, the word $S(x)$ is non-contractible.

Boundaries and overlappings: Given a word $w \in \mathscr{A}^*$, we define a *boundary* of w as a pair of consecutive positions $(p, p+1)$ in w. We will write simply p_w, the subindex denoting the corresponding word. By extension, we define $i(w) = 0_w$ and $f(w) = |w|_w$, the *initial* and *final* boundaries, respectively. Note that the boundaries of w have a natural linear order ($p_w \leqslant q_w$ iff $p \leqslant q$ as integers).

Given an equation $E = (w_1, w_2)$, an *overlapping* (of the boundaries of the left- and right-hand sides) of E is a linear order $\leqslant$ of the set of boundaries of w_1 and w_2 extending the natural orders of the boundaries of w_1 and w_2, such that $i(w_1) = i(w_2)$ and $f(w_1) = f(w_2)$ and possibly identifying some p_{w_1} and q_{w_2}.

Cuts and witnesses: Given an overlapping $\leqslant$ of $E = (w_1, w_2)$, a *cut* is a boundary j of w_2 (resp. w_1) such that $j \neq b$ for all boundaries b of w_1 (resp. w_2). Hence, a

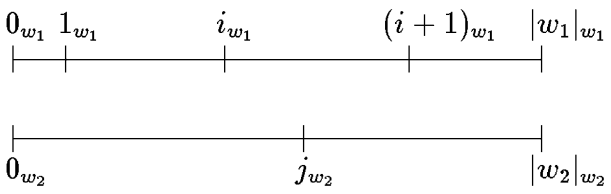

Fig. 1. The cut j_w.

cut determines at least three symbols of E, namely $w_2[j]$, $w_2[j+1]$ and $w_1[i+1]$, where i is such that $i_{w_1} < j_{w_2} < (i+1)_{w_1}$ in the linear order, see Fig. 1. The triple of symbols $(w_2[j], w_2[j+1], w_1[i])$ is called a *witness* of the cut. An overlapping is called *consistent* if $w_1[i+1]$ is a variable.

Observe that every overlapping gives rise to a system of equations $(E, \leqslant)$, which codifies the constraints given by $\leqslant$, by adding the corresponding equations and variables $x = x'y$ which the cuts determine. Also observe that every solution S of E determines a unique consistent overlapping, denoted $\leqslant_S$. Note finally that the cut j determines a boundary $(r, r+1)$ in $S(E)$; if $p \leqslant r < q$, we say that the factor $S(E)[p,q]$ of $S(E)$ *contains* the cut j.

Lemma 1. *Let E be an equation in FSI. Then E has a solution if and only if $(E, \leqslant)$ has a solution for some consistent overlapping $\leqslant$. There are no more than $|E|^{4|E|_V}$ consistent overlappings.*

Proof. Obviously, if for some consistent overlapping $\leqslant$, $(E, \leqslant)$ has a solution, then E has a solution. Conversely, if E has a solution S, consider the overlapping generated by S.

As for the bound, let $E = (w_1, w_2)$ and write v for $|E|_V$. First observe that if w_2 consists only of constants, then there are at most $|w_2|^v$ consistent overlappings. To get a consistent overlapping in the general case, first insert each initial and final boundary of each variable in w_2 in the linear order of the boundaries of w_1 (this can be done in at most $|E| + v$ ways). Then it remains to deal with the factors of w_2 in-between variables (hence consisting only of constants and of total length $\leqslant |E| - v$). Summing up, there are no more than $(|E| + v)^{2v}(|E| - v)^v \leqslant |E|^{4v}$ consistent overlappings. □

Lemma 2 (Compare Lemma 6, Rytter and Plandowski [16]). *Assume S is a minimal (w.r.t. length) solution of E. Then*

1. *For each factor $w = S(E)[i,j]$ with $|w| > 1$, there is an occurrence of w or w^{-1} which contains a cut of $(E, \leqslant_S)$.*
2. *For each letter $c = S(E)[i]$ of $S(E)$, there is an occurrence of c or c^{-1} in E.*

Proof. Let $1 \leqslant p \leqslant q \leqslant |S(E)|$. Suppose neither $w = S(E)[p,q]$ nor w^{-1} have occurrences in $S(E)$ which contain cuts. Consider the position p in $S(E)$ and its (S,E)-equivalence class P, and define for each variable x occurring in E,

$S'(x) =$ the subsequence of some image $S(x,i,j)$ of x consisting of all positions which are not in the set P. (i.e. "cut off" from $S(x,i,j)$ all the positions in P).

It is not difficult to see that S' is well defined, i.e., it does not depend on the particular image $S(x,i,j)$ of x chosen, and that $S'(w_1) = S'(w_2)$ (these facts follow from the definition of (S,E)-equivalence). Now, if P does not contain any images of the constants of E, it is easy to see that S' is a solution of the equation E. But $|S'(E)| < |S(E)|$, which is impossible because S was assumed to be minimal.

Hence, for each word $w = S[p,q]$, its first position must be in the same (S,E)-class of the position of the image of a constant c of E. If $p < q$, the right (resp. left) boundary of that constant is a cut in w (resp. w^{-1}) which is neither initial nor final (check definition of (S,E)-equivalence for $S(E)[p+1]$, etc.), and we are in case 1. If $p = q$ we are in case 2. □

Proposition 3. *For each non-contractible equation E there is a finite list of systems of equations $\Sigma_1, \ldots, \Sigma_k$ such that the following conditions hold*:
1. *E has a non-contractible solution if and only if one Σ_i has a solution.*
2. *$k \leqslant |E|^{8|E|_V}$.*
3. *There is $c > 0$ constant such that $|\Sigma_i| \leqslant c|E|$ and $|\Sigma_i|_V \leqslant c|E|_V$ for each $i = 1, \ldots, k$.*

Proof. Let $\leqslant$ be a consistent overlapping of E, and let

$$(x_1, y_1, z_1), \ldots, (x_r, y_r, z_r) \tag{2}$$

be a list of those witnesses of the cuts of $(E, \leqslant)$ for which at least one of the x_i, y_i is a variable. Let

$$D = \{(c,d) \in (C \cup C^{-1})^2 : \ c \neq d^{-1} \wedge d \neq c^{-1}\},$$

and define for each r-tuple $\langle (c_i, d_i) \rangle_i$, of pairs of D the system

$$\Sigma_{\langle (c_i,d_i)\rangle_i} = (E, \leqslant) \cup \{(x_i, x_i' c_i), (y_i, d_i y_i') : \ i = 1, \ldots, r\}.$$

Now, if S is a non-contractible solution of $(E, \leqslant)$ then S defines a solution of some Σ_i, namely the one defined by the r-tuple defined by the elements $(c_i, d_i) = (S(x_i)[|S(x_i)|], S(y_i)[1])$, for $i = 1, \ldots, r$. Note that because E and S are non-contractible, each (c_i, d_i) is in D.

On the other direction, suppose that S is a solution of some Σ_i. Then obviously S is a hand of $(E, \leqslant)$. We only need to prove that the $S(z)$ is non-contractible for all variables, z occurring in E. Suppose some z has a factor cc^{-1}, for $c \in C$. Then by Lemma 2 there is an occurrence of cc^{-1} (its converse is the same) which contains a cut of $(E, \leqslant)$. But because E is non-contractible, we must have that one of the terms in (2), say (x_j, y_j, z_j), witnesses this occurrence, hence $x_j = x_j' c$ and $y_j = c^{-1} y_j'$, which is impossible by the definition of the Σ_i's.

The bound in Condition 2 follows by simple counting: observe that $r \leqslant 2|E|_V$ and $|D| \leqslant |C|^{2r} \leqslant |E|^{4|E|_V}$, and the number k of systems is no larger than the number of overlappings times $|D|$. For the bounds in Condition 3 just sum the corresponding numbers of the new equations added. □

The following is an old observation of Hmelevskii [6] for free semigroups which can be extended to FSI:

Proposition 4. *For each system of equations Σ in FSI with generators C, there is an equation E in FSI with generators $C \cup c$, $c \notin (C \cup C^{-1})$, such that*
1. *S is a solution of E if and only if S is a solution of Σ.*
2. *$|E| \leqslant 4|\Sigma|$ and $|E|_V = |\Sigma|_V$.*
Moreover, if the equations in Σ are non-contractible, then E is non-contractible.

Proof. Let $(v_1, w_1), \ldots, (v_n, w_n)$ be the system of equations Σ. Define E as

$$(v_1 c v_2 c \cdots c v_n c v_1 c^{-1} v_2 c^{-1} \cdots c^{-1} v_n,\ w_1 c w_2 c \cdots c w_n c w_1 c^{-1} w_2 c^{-1} \cdots c^{-1} w_n).$$

Clearly, E is non-contractible because so was each equation (v_i, w_i), and c is a fresh letter. Also, if S is a solution of Σ, obviously it is a solution of E. Conversely, if S is a solution of E, then

$$|S(v_1 c v_2 c \cdots c v_n)| = |S(v_1 c^{-1} v_2 c^{-1} \cdots c^{-1} v_n)|,$$

hence

$$|S(v_1 c v_2 c \cdots c v_n)| = |S(w_1 c w_2 c \cdots c w_n)|,$$

and the same for the second pair of expressions with c^{-1} holds. Now it is easy to show that $S(v_i) = S(w_i)$ for all i: suppose not, for example $|S(v_1)| < |S(w_1)|$. Then $S(w_1)[|S(v_1)| + 1] = c$ and $S(w_1)[|S(v_1)| + 1] = c^{-1}$ are impossible. Then argue the same for the rest.

The bounds are simple calculations. □

The next theorem generalizes to FSI a key result in the study of solvability of word equations.

Theorem 5. *Let E be an equation in FSI. Then, the exponent of periodicity of a minimal solution of E is bounded by $2^{2^{\mathcal{O}(|E|)}}$.*

Proof. It is not worth reproducing here the ten-page proof in [8] because the changes needed to generalize it to FSI are minor. We will assume that the reader is familiar with the paper [8].

The proof there consists of two independent parts: (1) obtain from the word equation E a linear Diophantine equation, and (2) get a good bound for it. We will sketch how to do step (1) for FSI. The rest is completely identical.

First, let us sketch how the system of linear equations is obtained from a word equation E. Let S be a solution of E. Recall that a P-stable presentation of $S(x)$, for a variable x, has the form

$$S(x) = w_0 P^{\mu_1} w_1 P^{\mu_2} \ldots w_{n-1} P^{\mu_{n-1}} w_n.$$

From here, for a suitable P (which is the word that witnesses the exponent of periodicity of $S(E)$), a system of linear Diophantine equations $LD_P(E)$ is built, roughly speaking,

by replacing the μ_i by variables x_{μ_i} in the case of variables, plus some other pieces of data. Then it is proved that if S is a minimal solution of E, the solution $x_{\mu_i} = \mu_i$ is a minimal solution of $LD_P(E)$.

For the case of FSI, there are two key points to note. First, for the variables of the form x^{-1}, the solution $S(x^{-1})$ will have the following P^{-1}-stable presentation (same P, w_i, μ_i as before):

$$S(x^{-1}) = w_n^{-1}(P^{-1})^{\mu_{n-1}} w_{n-1}^{-1}(P^{-1})^{\mu_{n-2}} \cdots w_1^{-1}(P^{-1})^{\mu_1} w_0^{-1}.$$

Second, note that P^{-1} is a factor of PP if and only if P is a factor of $P^{-1}P^{-1}$. Call a repeated occurrence of P in w, say $w = uP^k v$, maximal, if P is neither the suffix of u nor a prefix of v. So it holds that maximal occurrences of P and P^{-1} in w either (1) do not overlap each other or (2) overlap almost completely (exponents will differ at most by 1).

In case (1), consider the system $LD_P(E') \cup LD_{P^{-1}}(E')$ (each one constructed exactly as in the case of word equations), where E' is the equation E where we consider the pairs of variables x^{-1}, x as independent for the sake of building the system of linear Diophantine equations. And, of course, the variables x_{μ_i} obtained from the same μ_i in $S(x)$ and $S(x^{-1})$ are the same.

In case (2), note that P-stable and P^{-1}-stable presentations for a variable x differ very little. So it is enough to consider $LD_P(E')$, taking care of using for the P-presentation of $S(x^{-1})$ the same set of Diophantine variables (adding 1 or -1 where it corresponds) used for the P-presentation of $S(x)$.

It must be proved then that if S is a minimal solution of the equation in FSI E, then the solution $x_{\mu_i} = \mu_i$ is a minimal solution of the corresponding system of linear Diophantine equations defined above. This can be proved easily with the help of Lemma 2.

Finally, as for the parameters of the system of Diophantine equations, observe that $|E'| = |E|$, hence, the only parameters that grow are the number of variables and equations, and by a factor of at most 2. So the asymptotic bound remains the same as for the case of E', which is $2^{\mathcal{O}(|E|)}$. □

The last result we will present concerning equations in FSI proves the simple and intuitive observation that every equation in free semigroups can be considered as an equation in FSI.

Proposition 6. *Let M be a free semigroup on the set of generators C, and N an FSI on the set of generators C, and E an equation in M. Then E is satisfiable in M if and only if it is satisfiable in N.*

Proof. An equation in FSI which does not contain $()^{-1}$ has a solution if and only if it has a solution which does not contain $()^{-1}$. So the codification of equations in free semigroups into FSI is straightforward: the same equation. □

We get immediately a lower bound for the problem of satisfiability of equations in FSI by using the corresponding result for the free semigroup case.

Corollary 7. *Satisfiability of equations in FSI is NP-hard.*

3. Reducing the problem of satisfiability of equations in free groups to satisfiability of equations in FSI

A *group* is an algebra with a binary associative operation (written as concatenation), a unary operation $()^{-1}$, and a constant 1, with axioms (1) plus

$$xx^{-1} = 1, \quad x^{-1}x = 1, \quad 1x = x1 = 1. \tag{3}$$

As in the case of FSI, it is not hard to see that the set of non-contractible words over $C \cup C^{-1}$ plus the empty word, and the operations of composition and the reverse suitably defined, is a free group with generators C.

Equations in free groups: The formal concept of *equation in free groups* is almost exactly the same as that for FSI; hence we will not repeat it here. The difference comes when speaking of solutions. A *solution* S of the equation E is (the unique group-homomorphism $S : \mathscr{A} \to (C \cup C^{-1})^*$ defined by) a map $S : V \to (C \cup C^{-1})^*$ extended by defining $S(c) = c$ for each $c \in C$ and $S(w^{-1}) = (S(w))^{-1}$, which satisfy $S(w_1) = S(w_2)$. Observe that the only difference with the case of SI is that now we possibly have 'simplifications' of subexpressions of the form ww^{-1} or $w^{-1}w$ to 1, i.e. the use of Eqs. (3).

Proposition 8 (Makanin [12, Lemma 1.1]). *For any non-contractible equation E in the free group G with generators C, we can construct a finite list $\Sigma_1, \ldots, \Sigma_k$ of systems of non-contractible equations in the FSI G' with generators C such that the following conditions are satisfied*:

1. *E has a non-contractible solution in G if and only if $k > 0$ and some system Σ_j has a non-contractible solution in G'.*
2. *There is $c > 0$ constant such that $|\Sigma_i| \leqslant |E| + c|E|_V^2$ and $|\Sigma_i|_V \leqslant c|E|_V^2$ for each $i = 1, \ldots, k$.*
3. *There is $c > 0$ constant such that $k \leqslant (|E|_V)^{c|E|_V^2}$.*

Proof. This is essentially the proof in [12] with the bounds improved. Let E be the equation

$$C_0 X_1 C_1 X_2 \cdots C_{v-1} X_v C_v = 1, \tag{4}$$

where C_i are non-contractible, $v = |E|_V$, and X_i are meta-variables representing the actual variables in E.

Let S be a non-contractible solution of E. By a known result (see [12, p. 486]), there is a set W of non-contractible words in the alphabet C, $|W| \leqslant 2v(2v+1)$, such that each C_i and $S(X_i)$ can be written as a concatenation of no more than $2v$ words in W, and after replacement Eq. (4) holds in the free group with generators W.

Let Z be a set of $2v(2v+1)$ fresh variables. Choose words $y_0, x_1, y_1, x_1, \ldots, x_v$, $y_v \in (Z \cup Z^{-1})^*$, each of length at most $2v$, non-contractible, and define the system of equations

1. $C_j = y_j$, $j = 0, \ldots, v$,
2. $X_j = x_j$, $j = 1, \ldots, v$.

Each such set of equations, for which Eq. (4) holds in the free group with generators Z when replacing C_i and X_i by the corresponding words in $(Z \cup Z^{-1})^*$, defines one system Σ_i.

It is clear from the result mentioned earlier that E has a solution if and only if there is some Σ_i which has a non-contractible solution. How many Σ_i are there? No more than $[(2v(2v+1))^{2v}]^{2v+1}$. □

Theorem 9. *For each equation E in a free group G with generators C there is a finite set Q of equations in a free semigroup with involution G' with generators $C \cup \{c_1, c_2\}$, $c_1, c_2 \notin C$, such that the following hold*:
1. *E is satisfiable in G if and only if one of the equations in Q is satisfiable in G'.*
2. *There is $c > 0$ constant, such that for each $E' \in Q$, it holds $|E'| \leqslant c|E|^2$.*
3. *Q can be generated non-deterministically in polynomial time.*

Proof. By Proposition 8, there is a list of systems of non-contractible equations $\Sigma_1, \ldots, \Sigma_k$ which are equivalent to E (w.r.t. non-contractible satisfiability). By Proposition 4, each such system Σ_j is equivalent (w.r.t. to satisfiability) to a non-contractible equation E'. Then, by Proposition 3, for each such non-contractible E', there is a system of equations (now without the restriction of non-contractibility) $\Sigma'_1, \ldots, \Sigma'_{k'}$ such that E' has a non-contractible solution if and only if one of the Σ'_j has a solution (not necessarily non-contractible). Finally, by Proposition 4, for each system Σ', we have an equation E'' which have the same solutions (if any) of Σ'. So we have a finite set of equations (the E'''s) with the property that E is satisfiable in G if and only if one of the E'' is satisfiable in G'.

The bounds in Condition 2 follow by easy calculations from the bounds in the corresponding results used above. Also, Condition 3 follows from a simple check. □

4. Conclusions

Our results show that solving equations in FSI comprises cases of free groups and free semigroups, the first with an exponential reduction (Theorem 9), and the latter with a linear reduction (Proposition 6). This suggest that FSI, due to its simplicity, is the 'appropriate' theory to study when seeking algorithms for solving equations in those theories.

After the results in [5], we conjecture that the computational complexity of the problem of satisfiability of equations in FSI is tightly linked to the corresponding problem for semigroups and groups, that is the problem is NP-complete.

Acknowledgements

Volker Diekert is thanked for useful comments. The author was supported by FONDAP, Matemáticas Aplicadas.

References

[1] V. Diekert, Makanin's Algorithm for Solving Word Equations with Regular Constraints, in: M. Lothaire (Ed.), Algebraic combinatorics on words, Cambridge Univ. Press, Cambridge, 2002.
[2] V. Diekert, C. Gutiérrez, Ch. Hagenah, The existential theory of equations with rational constraints in free groups is PSPACE-complete, in: Proc. STACS'2001, 18th Internat. Symp. on Theoretical Aspects of Computer Science, February 15–17, 2001, Dresden, Germany.
[3] C. Gutiérrez, Satisfiability of word equations with constants is in exponential space, in: Proc. 39th Annual IEEE Symposium on Foundations of Computer Science, FOCS'98, Palo Alto, 1998, IEEE Computer Soc. Press, pp. 112–119.
[4] C. Gutiérrez, Solving equations in strings: on Makanin's algorithm, in: Proc. Third Latin Amer. Symp. on Theoretical Informatics, LATIN'98, Campinas, Brazil, 1998, Lecture Notes in Computer Science, Vol. 1380, Springer, Berlin, pp. 358–373.
[5] C. Gutiérrez, Satisfiability of equations in free groups is in PSPACE, in: Proc. of the 32nd ACM Symposium on the Theory of Computing, STOC 2000, ACM Press, pp. 21–27.
[6] J.I. Hmelevskii, Equations in a free semigroup, Trudy Mat. Inst. Steklov 107 (1971). (English translation: Proc. Steklov Inst. Math. 107 (1971).)
[7] J. Jaffar, Minimal and complete word unification, J. ACM 37 (1) (1990) 47–85.
[8] A. Kościelski, L. Pacholski, Complexity of Makanin's algorithm, J. Assoc. Comput. Mach. 43 (1996) 670–684.
[9] A. Kościelski, L. Pacholski, Makanin's algorithm is not primitive recursive, Theoret. Comput. Sci. 191 (1998) 145–156.
[10] M. Lothaire, Combinatorics on Words, Cambridge Mathematical Texts, 1998, reprinted.
[11] G.S. Makanin, The problem of satisfiability of equations in a free semigroup, Mat. Sb. 103, 147–236 (in Russian). (English translation in Math. USSR Sb. 32, 129–198).
[12] G.S. Makanin, Equations in a free group, Iz. NA SSSR 46 (1982) 1199–1273 (in Russian). (English translation in Math USSR Iz. 21 (1983) 483–546.)
[13] W. Plandowski, Satisfiability of word equations with constants is in NEXPTIME, in: Proc. 31st Annual ACM Symposium on the Theory of Computing, STOC 99, ACM Press, pp. 721–725.
[14] W. Plandowski, Satisfiability of word equations with constants is in PSPACE, in: Proc. 40th Annual Symposium on Foundations of Computer Science, FOCS 99, IEEE Computer Society Press, 1999, pp. 495–500.
[15] A.A. Razborov, On systems of equations in a free group, Izv. AN SSSR 48 (1984) 779–832 (in Russian). (English translation in Math. USSR Izv. 25 (1985) 115–162.)
[16] W. Rytter, W. Plandowski, Applications of Lempel–Ziv encodings to the solution of word equations, in: K.G. Larsen et al. (Eds.), Proc. of the 25th ICALP, Lect. Notes in Comp. Science, Springer, 1998, pp. 731–742.
[17] K. Schulz, Word unification and transformation of generalized equations, J. Automat. Reason. 11 (1993) 149–184.