Solving Word Equations Modulo Partial Commutations

Volker Diekert[1], Yuri Matiyasevich[2*], and Anca Muscholl[1]

[1] Institut für Informatik, Universität Stuttgart,
Breitwiesenstr. 20-22, 70565 Stuttgart, Germany
[2] Steklov Institute of Mathematics at St.Petersburg
Fontanka 27, St. Petersburg, 191011 Russia

Abstract. It is shown that it is decidable whether an equation over a free partially commutative monoid has a solution. We give a proof of this result using normal forms. Our method is a direct reduction of a trace equation system to a word equation system with regular constraints. Hereby we use the extension of Makanin's theorem on the decidability of word equations to word equations with regular constraints, which is due to Schulz.

Keywords: word equations, partial commutations, traces, word equations with regular constraints

1 Introduction

Solving equations is a central topic in various fields of computer science, especially concerning unification, as required by automated theorem proving or logic programming. A famous result of Makanin [15] states that there exists an algorithm deciding for a given equation $L = R$, where $L, R \in (\Omega \cup \Sigma)^*$ contain both unknowns from Ω and constants from Σ, whether an assignment $\sigma: \Omega \to \Sigma^*$ exists satisfying $\sigma(L) = \sigma(R)$. In a more general setting it is known that the existential theory of equations over free monoids is decidable, i.e., given an existentially quantified, closed first-order formula S over atomic predicates of the form $L = R$ and $L \neq R$, it is decidable whether S is valid over a given free monoid. Moreover, adding regular constraints, i.e., atomic predicates of the form $x \in C$, where C is a regular language, preserves decidability, as it was shown by Schulz [22].

In this paper we prove the generalization of Makanin's result to equations over words modulo a given partial commutation, i.e. we generalize it to trace monoids. These monoids were originally studied in combinatorics [5]. They became meaningful for computer science in concurrency theory, where they were introduced by Mazurkiewicz [20] in connection with the semantics of labelled

* The main part of this work was done during a stay of the author at the University of Stuttgart.

Petri nets and by Keller for investigating parallel program schemata [14]. For an overview of trace theory and related topics we refer to "The Book of Traces" [9].

Most results obtained so far in the area of equations on traces were restricted to equations without constants, see [6,10]. The decidability of the solvability of equations with constants was stated as an important open question. In this paper we present a positive answer to this question. The original proof of the second author [17], which was presented at LFCS'97 [18], goes by induction over the size of the alphabet. A direct proof based on lexicographical normal forms was presented at ICALP'97 [8].

The paper is structured as follows. Section 2 introduces basic notions. Section 3 contains the outline of the proof idea and the statements of the main results. In Section 4 we consider normal forms for trace monoids. We exhibit a new result concerning the product of lexicographical normal forms in Section 4.1, whereas Section 4.2 establishes a link between the two decision procedures for solving trace equations given in [8,18], by introducing regular normal forms. In Section 5 we conclude the proofs of the main results and discuss some extensions. Section 6 considers mixed systems of word and trace equations and finally we give an application to the parallel complexity of computing lexicographical normal forms in Section 7.

2 Notations and Preliminaries

An *independence alphabet* is a pair (Σ, I), where Σ is a finite alphabet of cardinality $n \geq 1$ and $I \subseteq \Sigma \times \Sigma$ is an irreflexive and symmetric relation, called *independence relation* (or *commutation relation*).

With a given independence alphabet (Σ, I) we associate the *trace monoid* $\mathbb{M}(\Sigma, I)$. This is the quotient monoid $\Sigma^*/\equiv_I$, where $\equiv_I$ denotes the congruence being the equivalence relation generated by the set $\{uabv = ubav \mid (a,b) \in I,\, u,v \in \Sigma^*\}$; an element $t \in \mathbb{M}(\Sigma, I)$ is called a *trace*, the length $|t|$ of a trace t is given by the length of any representing word. By $\mathrm{alph}(t)$ we denote the alphabet of a trace t, being the set of letters occurring in t.

By 1 we denote both the empty word and the empty trace. Words $v, w \in \Sigma^*$ are called *independent* if $\mathrm{alph}(v) \times \mathrm{alph}(w) \subseteq I$. In this case we simply write $(v,w) \in I$ or $v \in I(w)$, where $I(w)$ for $w \in \Sigma^*$ is a shorthand for $\{a \in \Sigma \mid \{a\} \times \mathrm{alph}(w) \subseteq I\}$.

The *initial alphabet* of $w \in \Sigma^*$ is the set $\mathrm{init}(w) = \{a \in \Sigma \mid \exists w' \in \Sigma^* \text{ with } w \equiv_I aw'\}$.

3 Trace Equation Systems

Definition 1. Let (Σ, I) denote an independence alphabet and let Ω be a finite set of unknowns (variables), $\Sigma \cap \Omega = \emptyset$. A trace equation with unknowns Ω over (Σ, I) has the form $L \equiv R$, with $L, R \in (\Sigma \cup \Omega)^*$. An assignment for $L \equiv R$ is a mapping $\sigma : \Omega \to \Sigma^*$. Let $\sigma : (\Sigma \cup \Omega)^* \to \Sigma^*$ denote also the homomorphic extension of σ given by $\sigma|_\Sigma = \mathrm{id}_\Sigma$. An assignment $\sigma : \Omega \to \Sigma^*$ is called a solution for $L \equiv R$ if $\sigma(L) \equiv_I \sigma(R)$, i.e. $\sigma(L)$ and $\sigma(R)$ are the same trace in $\mathbb{M}(\Sigma, I)$. A system T of trace equations is a finite conjunction

$$L_1 \equiv R_1 \ \& \ \cdots \ \& \ L_k \equiv R_k$$

of trace equations. A mapping $\sigma : \Omega \to \Sigma^*$ is a solution for T if we have $\sigma(L_i) \equiv_I \sigma(R_i)$ for all i, $1 \le i \le k$.

Notation. In the special case of words, i.e. $I = \emptyset$, we denote equations simply as $L = R$, with $L, R \in (\Sigma \cup \Omega)^*$.

Example 2. Let $\Sigma = \{a, b\}$. If $(a, b) \in I$ then for the trace equation $ax \equiv xa$ over (Σ, I) every assignment is a solution. The word equation $ax = xa$ has just solutions assigning x some word from a^*.

Lemma 3. *For every system of trace equations T with unknowns Ω an equivalent system T' of trace equations with unknowns Ω' can be effectively determined such that every equation in T' has the form $xy \equiv z$, with $x, y, z \in \Omega \cup \Sigma$. Here we mean by equivalence that the sets of solutions is bijectively related in such a way that a solution σ for T is mapped to a solution σ' for T' if and only if $\sigma|_{\Omega\cap\Omega'} = \sigma'|_{\Omega\cap\Omega'}$.*

Proof. Suppose that $x_1 \cdots x_l \equiv x_{l+1} \cdots x_r$, with $x_i \in \Omega \cup \Sigma$ is an equation in T, $0 \le l < r$. We may assume that $l > 0$ (formally, an equation $1 \equiv x_{l+1} \cdots x_r$ can be eliminated after removing all occurrences of x_i, $l < i \le r$). For such an equation we introduce new variables $y_1, \ldots, y_r$, together with new equations

$$\begin{array}{rcr} x_1 \equiv y_1, & & x_{l+1} \equiv y_{l+1} \\ y_1 x_2 \equiv y_2, & & y_{l+1} x_{l+2} \equiv y_{l+2}, \\ \vdots & & \vdots \\ y_{l-1} x_l \equiv y_l, & & y_{r-1} x_r \equiv y_r, \\ & y_l \equiv y_r & \end{array}$$

Obviously, the set of solutions for $x_1 \cdots x_l \equiv x_{l+1} \cdots x_r$ is in canonical bijection with the set of solutions of the above system. Finally, trivial equations $x \equiv y$ can be eliminated by replacing every occurrence of y by x.

We will show the decidability of trace equations by transforming them into word equations with simple additional constraints of the form $(x, y) \in I$, where $x, y \in \Omega$ are unknowns. Constraints of the form $(x, y) \in I$ are denoted as *commutation constraints.* Their meaning is explained in the following definition:

Definition 4. A system S of *word equations* with unknowns Ω and *commutation constraints* over (Σ, I) is given as a pair consisting of a set W of word equations over Σ and Ω, and a set C of constraints of the form $(x, y) \in I$, with $x, y \in \Omega$.

A mapping $\sigma : \Omega \to \Sigma^*$ is a *solution* for S if

i) $\sigma(L) = \sigma(R)$ is satisfied for all equations $L = R$ from W and
ii) $\mathrm{alph}(\sigma(x)) \times \mathrm{alph}(\sigma(y)) \subseteq I$ is satisfied for all commutation constraints $(x, y) \in I$ from C.

For a natural number α let $\mathcal{W}_\alpha(x, y, z)$ denote the set of all systems of word equations with commutation constraints over (Σ, I) of the form

$$\begin{aligned}
& x = x_{i_1} \cdots x_{i_\alpha} \;\; \& \\
& y = y_{j_1} \cdots y_{j_\alpha} \;\; \& \\
& z = x_1 y_1 \cdots x_\alpha y_\alpha \;\; \& \\
& (x_{i_p}, x_{i_q}) \in I \text{ for all } p, q \text{ such that } (i_p - i_q)(p - q) < 0 \;\; \& \\
& (y_{j_p}, y_{j_q}) \in I \text{ for all } p, q \text{ such that } (j_p - j_q)(p - q) < 0 \;\; \& \\
& (x_i, y_j) \in I \text{ for all } i > j \,,
\end{aligned}$$

where $(i_1, \ldots, i_\alpha)$ and $(j_1, \ldots, j_\alpha)$ are permutations of $(1, \ldots, \alpha)$ and x_i, y_j are new variables.

Remark 5. Consider any system S from $\mathcal{W}_\alpha(x, y, z)$ for some α. Obviously, if σ is a solution for S and $u = \sigma(x)$, $v = \sigma(y)$ and $w = \sigma(z)$ then $uv \equiv_I w$ due to the commutation constraints.

The reduction from trace equations to word equations with commutation constraints is based on the idea of using normal forms for traces and preserving solutions in normal form. Namely, a suitable normal form allows us to state the converse of the above remark.

First recall the definition of normal forms:

Definition 6. Let (Σ, I) be an independence alphabet. Then $\mathcal{N} \subseteq \Sigma^*$ is called a set of normal forms for $\mathbb{M}(\Sigma, I)$ if we have

$$|\mathcal{N} \cap \{v \mid v \equiv_I u\}| = 1$$

for every $u \in \Sigma^*$.

We denote below by $\nu_{\mathcal{N}}(u)$ for $u \in \Sigma^*$ the unique word $v \in \mathcal{N}$ such that $u \equiv_I v$, that is, $\nu_{\mathcal{N}}$ is a mapping $\nu_{\mathcal{N}} : \Sigma^* \to \mathcal{N}$.

The main reduction step will rely on a suitable choice of a set of normal forms $\mathcal{N}$, enabling us to state the converse of Rem. 5 in the following way:

Theorem 7. *For every trace monoid $\mathbb{M}(\Sigma, I)$ there is an effectively computable number α such that for a suitable set of normal forms $\mathcal{N} \subseteq \Sigma^*$ for $\mathbb{M}(\Sigma, I)$ the following holds: For every words u, v, w from $\mathcal{N}$ such that $uv \equiv_I w$ there is a system S from $\mathcal{W}_\alpha(x, y, z)$ having a solution σ such that $u = \sigma(x)$, $v = \sigma(y)$, and $w = \sigma(z)$.*

Remark 8. We cannot expect to obtain for every trace equation $xy \equiv z$ a finite set $\mathcal{W}(x, y, z)$ of systems of word equations with commutation constraints over unknowns $\Omega \supseteq \{x, y, z\}$ such that the following stronger requirement (not using normal forms as in Thm. 7) is satisfied:

For every $u, v, w \in \Sigma^*$ with $uv \equiv_I w$ there exists a system $S \in \mathcal{W}(x, y, z)$ and a solution σ for S such that $\sigma(x) = u$, $\sigma(y) = v$, and $\sigma(z) = w$. Moreover, for every $S \in \mathcal{W}(x, y, z)$ and every solution σ for S, $\sigma(x)\sigma(y) \equiv_I \sigma(z)$.

To see this, consider the independence alphabet (Σ, I), $\Sigma = \{a, b\}$ with $(a, b) \in I$, together with the following system S of trace equations with commutation constraints:

$$xy \equiv z \;\&\; x \equiv ax' \;\&\; y \equiv by' \;\&\; (x, y) \in I\,.$$

Every solution σ for S has the form $\sigma(x) = a^i$, $\sigma(y) = b^j$, $\sigma(z) = w$ with $w \in \{a, b\}^*$, $|w|_a = i > 0$, $|w|_b = j > 0$. Note also that a system of *word* equations with commutation constraints over (Σ, I) as above can be easily transformed into a system without constraints, since e.g. $x \in a^*$ is equivalent to the word equation $ax = xa$.

Assume that a finite collection $\mathcal{W}$ of systems of word equations exists for each trace equation in the above system, such that the stronger requirement above is satisfied. We obtain a finite set $\mathcal{W}'$ of systems of word equations such that the language $\{a^i b^j aw \mid i > 0, j > 0, w \in \{a, b\}^*, |w|_a + 1 = i, |w|_b = j\}$ is expressible as the set of all $\sigma(t)$, where σ is a solution for some system $S \in \mathcal{W}'$ and t is a fixed unknown. But this contradicts Thm. 11 from [13]. A direct proof of this claim will also follow from the results of Section 6.

Thm. 7 will yield directly the decidability of systems of trace equations. First, note that commutation constraints $(x, y) \in I$ can be transformed into regular constraints over the alphabet $\Sigma \cup \{d\}$, where $d \notin \Sigma$. For this we introduce new unknowns z and new equations $z = xdy$, together with constraints $z \in C$, where $C = \{udv \mid u, v \in \Sigma^*, (u, v) \in I\}$. Thm. 7 has the following consequence for systems of trace equations (recall that $\nu_{\mathcal{N}} : \Sigma^* \to \mathcal{N}$ denotes the canonical mapping w.r.t. the set of normal forms $\mathcal{N}$). By Lem. 3 we start with a system of trace equations of the form $xy \equiv z$, $x, y, z \in \Omega \cup \Sigma$. Then we replace each such trace equation by a disjunction over all systems of word equations with commutation constraints from $\mathcal{W}_\alpha(x, y, z)$. By distributivity we obtain the finite set $\mathcal{W}$ of systems of word equations with regular constraints in the corollary below:

Corollary 9. *Let T be a system of trace equations with unknowns Ω over (Σ, I), and let $d \notin \Sigma$. Then a set of unknowns $\Omega' \supseteq \Omega$ and a finite set $\mathcal{W}$ of systems of word equations with unknowns Ω' and regular constraints over $\Sigma \cup \{d\}$ can be effectively constructed such that we have:*

i) For every solution σ for T there is some $S \in \mathcal{W}$ such that the mapping $\nu_{\mathcal{N}} \circ \sigma$ can be extended to a solution for S.
ii) For every $S \in \mathcal{W}$ and every solution σ for S, $\sigma|_\Omega$ is a solution for T.

Using encodings (see Prop. 32 in Sect. 5) we can avoid extending the alphabet for the regular constraints, thus obtaining:

Corollary 10. *The solvability of a system of trace equations over (Σ, I) is reducible to the solvability of some system of word equations with regular constraints over Σ from an effectively computable finite set of such equation systems.*

Moreover, a system of word equations $L_1 = R_1 \,\&\, \cdots \,\&\, L_k = R_k$ can be transformed by classical techniques (see e.g. [7]) into a single word equation. Hence, a system of trace equations over (Σ, I) can be reduced to a disjunction of word equations with regular constraints over Σ. The solvability of each of these word equations is decidable by Schulz' generalization of Makanin's theorem, [15,22]. Hence, we obtain:

Corollary 11. *It is decidable whether a system of trace equations has a solution.*

4 Normal Forms

In order to show Thm. 7 we need a suitable set $\mathcal{N}$ of normal forms for $\mathbb{M}(\Sigma, I)$ such that we can compute some value $\alpha(\mathcal{N})$ which suffices for showing the following: For all words $u, v, w \in \mathcal{N}$ in normal form we can decompose u, v, w as

$$\begin{aligned} w &= u_1 v_1 \cdots u_\alpha v_\alpha, \\ u &= u_{i_1} \cdots u_{i_\alpha}, \\ v &= v_{j_1} \cdots v_{j_\alpha}, \end{aligned} \tag{1}$$

for some $\alpha \leq \alpha(\mathcal{N})$, permutations $(i_1, \ldots, i_\alpha)$, $(j_1, \ldots, j_\alpha)$ of $(1, \ldots, \alpha)$ and words $u_i, v_j \in \mathcal{N}$ satisfying

$$\begin{aligned} &(u_{i_p}, u_{i_q}) \in I \text{ for all } p, q \text{ with } (i_p - i_q)(p - q) < 0, \\ &(v_{j_p}, v_{j_q}) \in I \text{ for all } p, q \text{ with } (j_p - j_q)(p - q) < 0, \\ &(u_i, v_j) \in I \text{ for all } i > j. \end{aligned} \tag{2}$$

In the following subsections we will show how to obtain $\alpha(\mathcal{N})$ for two natural sets of normal forms, the lexicographical normal form and the priority normal form.

4.1 Lexicographical normal form

Let $<$ denote a fixed total order on the alphabet Σ. A word $v \in \Sigma^*$ is in *lexicographical normal form* (w.r.t. I and $<$) if v is the least word in $\{w \mid v \equiv_I w\}$ w.r.t. the lexicographical order. Let LNF denote the set of lexicographical normal forms, i.e., $\mathrm{LNF} \subseteq \Sigma^*$ is the set of minimal representatives for $\mathbb{M}(\Sigma, I)$. For $v \in \Sigma^*$ we denote by $\ell nf(v)$ the lexicographical normal form of v, that is, $\ell nf : \Sigma^* \to \mathrm{LNF}$. When appropriate, we identify Σ with the set $\{1, \ldots, n\}$.

The lexicographical normal form is connected to a variant of sorting studied by Anisimov and Knuth [3]. They gave an alternative characterization of lexicographical normal forms showing that LNF is actually a regular, even star-free language:

Proposition 12 ([3]). *Let Σ be totally ordered by $<$. Then a word $v \in \Sigma^*$ is in lexicographical normal form (w.r.t. I and $<$) if and only for every factor aub of v with $a, b \in \Sigma$, $u \in \Sigma^*$ and $(au, b) \in I$ we have $a < b$.*

Definition 13. Let Σ be totally ordered by $<$. For $\emptyset \neq A \subseteq \Sigma$ let the *height* $h(A)$ be $h(A) = \max\{a \mid a \in A\}$. Let also $h(\emptyset) = 0$. (Thus, $h(A) \in \{0, \ldots, n\}$.)
The *height* $h(v)$ of a word $v \in \Sigma^*$ is defined as $h(v) = h(\mathrm{alph}(v))$.

The main arguments for showing that lexicographical normal forms satisfy the above requirements (Eqs. (1), (2)) are given in the lemma below. There are two important aspects in this lemma. The first one concerns the fact that the value of m below is uniformly bounded, whereas the second one is that the height of the left factor will decrease strictly.

Lemma 14. *Let $u, v, w \in \mathrm{LNF}$ be words in lexicographical normal form such that $uv \equiv_I w$. Let $h = h(u)$ be the height of u and suppose that $h > 0$. Then there exist an integer m, $1 \leq m \leq \frac{(n-1)(h-1)}{2} + 1$, and words $u_1, \ldots, u_m$, $v_1, \ldots, v_m \in \mathrm{LNF}$ in lexicographical normal form such that the following conditions are satisfied:*

$$
\begin{aligned}
&u = u_1 \cdots u_m \quad \textit{(word equality)},\\
&v \equiv_I v_1 \cdots v_m \quad \textit{(trace equality)},\\
&w = u_1 v_1 \cdots u_m v_m \quad \textit{(word equality)},\\
&u_i \neq 1 \ \textit{ for all } \ 1 < i \leq m\,,\\
&v_j \neq 1 \ \textit{ for all } \ 1 \leq j < m\,,\\
&v_j \in I(u_{j+1} \cdots u_m) \textit{ for all } \ 1 \leq j < m\,,\\
&h(v_j) < h \textit{ for all } \ 1 \leq j < m\,.
\end{aligned}
$$

Remark 15. Before giving the proof of Lem. 14, let us note that the trace equality $uv \equiv_I w$ above cannot be replaced by word equalities of type $u = u_1 \cdots u_m$, $v = v_1 \cdots v_m$, $w = u_1 v_1 \cdots u_m v_m$. For example, consider $\mathbb{M}(\Sigma, I) = \{a, b, c\}^* / \{ab = ba, bc = cb\}$ and $u = c$, $v = ab$. Then the lexicographical normal form of uv is $w = bca$.

Proof of Lem. 14. We have $uv \equiv_I w$ with $u, v, w \in \mathrm{LNF}$ and $h = h(u) > 0$. Consider the decomposition of w, $w = u_1 v_1 \cdots u_m v_m$, where $m \geq 1$ is minimal such that $u \equiv_I u_1 \cdots u_m$, $v \equiv_I v_1 \cdots v_m$, and $v_j \in I(u_{j+1} \cdots u_m)$ for all j, $1 \leq j < m$. Since m is minimal, we have $u_i \neq 1$ and $v_j \neq 1$ for all $1 < i \leq m$, $1 \leq j < m$. Moreover, the words u_i, v_j are in lexicographical normal form.

Let us first show that $u = u_1 \cdots u_m$. Assume that atb is a factor of $u_1 \cdots u_m$ with $a, b \in \Sigma$, $t \in \Sigma^*$ and $b \in I(at)$. If atb is a factor of some u_i, then $a < b$ follows by Prop. 12 and we are done. Otherwise let $i < j$ be such that $u_i \in \Sigma^* at'$, $u_j \in t'' b \Sigma^*$ and $t = t' u_{i+1} \cdots u_{j-1} t''$. Since $v_k \in I(u_j)$ for $k < j$ we obtain $b \in I(at' u_{i+1} v_{i+1} \cdots u_{j-1} v_{j-1} t'')$, hence $a < b$ due to w being in lexicographical normal form. Thus $u_1 \cdots u_m$ is in lexicographical normal form, again by Prop. 12, and it follows that $u = u_1 \cdots u_m$.

Suppose that $1 \leq j < m$ and let b denote the first letter of u_{j+1}. Let $a \in \mathrm{alph}(v_j)$, i.e. $v_j = tat'$ for some words t, t'. Then $at'b$ is a factor of $w \in \mathrm{LNF}$

satisfying $b \in I(at')$, thus we have $a < b$. Therefore $h(v_j) < h(b) \leq h$ for every $1 \leq j < m$.

Finally, assume by contradiction that $m > (n-1)(h-1)/2 + 1$. Let b_i, a_j denote the first letter of u_i, v_j respectively, $1 < i \leq m$, $1 \leq j < m$. Consider the chain of alphabets $I(u_2 \cdots u_m) \subseteq I(u_3 \cdots u_m) \subseteq \cdots \subseteq I(u_m)$. Note that we have $I(u_2 \cdots u_m) \neq \emptyset$ due to $v_1 \neq 1$, and also $I(u_m) \neq \Sigma$ due to $u_m \neq 1$. Therefore by the pigeon-hole principle there exist some indices $1 \leq i, j < m$ with $j - i \geq (h-1)/2$ satisfying $I(u_{i+1} \cdots u_m) = I(u_{j+1} \cdots u_m)$. Consider the factor $v_i u_{i+1} v_{i+1} \cdots v_j u_{j+1}$ of w. Note that we have $(v_k, u_l) \in I$ for every k, l such that $i \leq k, l-1 \leq j$, since $v_k \in I(u_{k+1} \cdots u_m) = I(u_{i+1} \cdots u_m)$. Therefore, $w \in \mathrm{LNF}$ implies $a_i < b_{i+1} < a_{i+1} < \cdots < a_j < b_{j+1}$ and we obtain $h(u) \geq h(b_{j+1}) \geq 2(j-i+1) > h$, a contradiction.

Lem. 14 can be applied inductively in order to satisfy the conditions given in Eqs. (1), (2). Recall that $n = |\Sigma|$.

Proposition 16. *Let $u, v, w \in \mathrm{LNF}$ be words in lexicographical normal form such that $uv \equiv_I w$. Then there exists a positive integer $\alpha \leq n^n n!$ such that*

$$\begin{aligned} u &= u_1 \cdots u_\alpha \,, \\ v &= v_{j_1} \cdots v_{j_\alpha} \,, \\ w &= u_1 v_1 \cdots u_\alpha v_\alpha \,, \end{aligned}$$

for some permutation $(j_1, \ldots, j_\alpha)$ of $(1, \ldots, \alpha)$ and words $u_i, v_j \in \mathrm{LNF}$ satisfying

$$\begin{aligned} &(u_i, v_j) \in I \text{ for all } i > j \text{ and} \\ &(v_{j_p}, v_{j_q}) \in I \text{ for all } p, q \text{ with } (j_p - j_q)(p-q) < 0\,. \end{aligned}$$

Proof. Let $h = h(u)$ denote the height of u. The proof goes by induction on h. Let $\alpha = \alpha(h)$ be minimal such that

$$\begin{aligned} u &= u_1 \cdots u_\alpha \,, \\ v &= v_{j_1} \cdots v_{j_\alpha} \,, \\ w &= u_1 v_1 \cdots u_\alpha v_\alpha \,, \end{aligned}$$

for some words u_i, v_j satisfying the commutation conditions of the proposition. Clearly, $\alpha(h)$ is trivially bounded by the length of w. For $h = 0$ we have $u = 1$, thus $\alpha(0) = 1$, which satisfies the proposition.

Let $h \geq 1$. By Lem. 14 there exist an integer $m \leq (n-1)(h-1)/2 + 1$ and words $u_1, \ldots, u_m$, $v_1, \ldots, v_m$ in lexicographical normal form satisfying

$$\begin{aligned} u &= u_1 \cdots u_m \,, \\ v &\equiv_I v_1 \cdots v_m \,, \\ w &= u_1 v_1 \cdots u_m v_m \,, \\ u_i \neq 1,\ v_j &\neq 1 \ \text{ for } 1 < i \leq m,\ 1 \leq j < m \,, \\ v_j \in I(u_{j+1} \cdots u_m) &\text{ and } h(v_j) < h \text{ for } 1 \leq j < m \,. \end{aligned} \tag{3}$$

If $h = 1$, then $m = 1$ in (3), so we can take $\alpha(1) = 1$, since $v = v_1 \in$ LNF, which satisfies the claim. Hence let $h, m \geq 2$. Let $\bar{v}_1 = v_1$ and $\bar{v}_i = \ell nf(\bar{v}_{i-1} v_i)$ for $i = 2, \ldots, m$. Clearly, $\bar{v}_m = v$, $h(\bar{v}_i) < h$ for $1 \leq i < m$ and

$$\bar{v}_{i-1} v_i \equiv_I \bar{v}_i, \quad \text{for } 1 < i \leq m. \tag{4}$$

We can apply the induction hypothesis to each of the $(m-1)$ equivalences (4). We obtain that some $\beta \leq \alpha(h-1)$ and some (not necessarily non-empty) words $s_{1,i}, \ldots, s_{\beta,i}$, $t_{1,i}, \ldots, t_{\beta,i}$ exist satisfying:

$$\begin{aligned} \bar{v}_i &= s_{1,i} t_{1,i} \cdots s_{\beta,i} t_{\beta,i} \,, \\ \bar{v}_{i-1} &= s_{1,i} \cdots s_{\beta,i} \,, \\ v_i &= t_{j_1,i} \cdots t_{j_\beta,i} \,. \end{aligned}$$

(By induction, the words $s_{k,i}, t_{j,i}$ also satisfy commutation conditions as required in the proposition. Moreover, $(j_1, \ldots, j_\beta)$ is a permutation of $(1, \ldots, \beta)$.)

Using the identities

$$s_{1,i} \cdots s_{\beta,i} = s_{1,i-1} t_{1,i-1} \cdots s_{\beta,i-1} t_{\beta,i-1}$$

for $i = m, \ldots, 2$ and adequately splitting the factors $t_{1,i-1}, \ldots, t_{\beta,i-1}$ we obtain by renaming the factors $t_{j,i}$ into $v_1, \ldots, v_\alpha$, resp. renaming the factors $s_{j,i}$ into u_i:

$$\begin{aligned} u &= u_1 \cdots u_\alpha \,, \\ v &= v_{j_1} \cdots v_{j_\alpha} \,, \\ w &= u_1 v_1 \cdots u_\alpha v_\alpha \,, \end{aligned}$$

for $v = \bar{v}_m$ and u_i, v_j satisfying the commutation conditions and for some α such that

$$\alpha \leq (m-1) \cdot 2 \cdot \beta \leq 2(m-1)\alpha(h-1)\,.$$

This yields

$$\alpha(h) \leq (n-1)(h-1)\alpha(h-1) \leq (n-1)^{h-1}(h-1)! \leq n^n n!\,,$$

which concludes the proof.

4.2 Regular normal forms

The aim of this section is to establish a link between the proof using lexicographical normal forms of the previous section and the original proof of the second author. The proof of Cor. 9 given in [18] goes by induction over the size of the independence relation I. However, the proof can be also seen as using another normal form for traces, which we denote below as *priority normal form*. The natural question arising is which properties of the lexicographical resp. priority normal form are important for obtaining a decomposition as in Prop. 16.

Remark 17. Note that requiring that a set of normal forms $\mathcal{N} \subseteq \Sigma^*$ is regular does not suffice for obtaining an analogue of Prop. 16. Consider e.g. the independence alphabet $(\{a,b\}, I)$ with $(a,b) \in I$. The set $\mathcal{N} = (ab)^*(a^* + b^*)$ is a set of normal forms for $\mathbb{M}(\Sigma, I)$. But with $u = (ab)^m b^m$, $v = (ab)^m a^m$ the normal form of uv is $\nu(uv) = (ab)^{3m}$ which cannot be represented as in Prop. 16 for any large enough value of m.

We introduce in this section a class of regular normal forms which we call right-factor resp. left-factor normal forms. We show that these notions generalize the lexicographical normal form, as well as the priority normal form used in [18]. Moreover, right-factor resp. left-factor normal forms suffice in order to obtain an analogue of Prop. 16. As a byproduct we obtain an alternative characterization of the priority normal form.

Definition 18 ([21]). A monoid S is called *ordered* if there exists some total order $<$ such that $s \leq t$ implies $xsy \leq xty$ for all $x, y \in S$.

For a homomorphism $h : \Sigma^* \to S$, a monoid element $s \in S$ and a subalphabet $A \subseteq \Sigma$, we denote by $h_A^{-1}(s)$ the set $h^{-1}(s) \cap \{w \mid \mathrm{alph}(w) = A\}$ of words with alphabet A which are mapped to s.

Definition 19. Let $\mathcal{N} \subseteq \Sigma^*$ be a set of normal forms for $\mathbb{M}(\Sigma, I)$. Then $\mathcal{N}$ is called a set of *right-factor* normal forms if there exist a finite ordered monoid $(S, <)$ and a homomorphism $h : \Sigma^* \to S$ such that

i) $\mathcal{N} = \Sigma^* \ \setminus \bigcup_{\emptyset \neq A \times B \subseteq I,\ t \leq s} \Sigma^* \ h_A^{-1}(s) h_B^{-1}(t) \ \Sigma^*$

ii) $(xyz \in \mathcal{N} \ \wedge \ (y,z) \in I) \quad \Longrightarrow \quad xz \in \mathcal{N}$.

Remark 20. $\mathcal{N}$ is called a set of *left-factor* normal forms if Cond. (ii) of Def. 19 is replaced by

$$(xyz \in \mathcal{N} \ \wedge \ (x,y) \in I) \quad \Longrightarrow \quad xz \in \mathcal{N}.$$

We will show at the end of the section that right-factor (resp. left-factor) normal forms can be used for reducing trace equations to word equations. But let us first illustrate this notion of normal forms by two natural examples.

For lexicographical normal forms we obtain immediately by Prop. 12 that LNF is a set of right-factor normal forms:

Proposition 21. *Let Σ be totally ordered by $<$ and let (Σ, I) be an independence alphabet. Then* LNF *is a set of right-factor normal forms for $\mathbb{M}(\Sigma, I)$.*

Proof. Consider the monoid $S = (\Sigma \dot\cup \{1\}, \cdot, 1)$ with $a \cdot b = \max(a,b)$ for all $a, b \in \Sigma$. It is straightforward to verify that together with $1 \leq a$ for all $a \in \Sigma$ we obtain an ordered monoid, $S = (S, <)$. Let $h : \Sigma^* \to S$ yurisugg Let homomorphism $h : \Sigma^* \to S$ be given by $h(a) = a$, then $h(1) = 1$ and $h(w) = \max(\mathrm{alph}(w))$ for all $w \neq 1$.

We show in the following that

$$\mathrm{LNF} = \Sigma^* \ \setminus \bigcup_{\emptyset \neq A\times B \subseteq I,\, t \leq s} \Sigma^* \ h_A^{-1}(s) h_B^{-1}(t) \ \Sigma^* \, . \tag{5}$$

Suppose first that $xuvy \in \mathrm{LNF}$, $x,u,v,y \in \Sigma^*$, $\mathrm{alph}(u) = A \neq \emptyset$, $\mathrm{alph}(v) = B \neq \emptyset$, $A \times B \subseteq I$. Let $a = \max(A)$, $b = \max(B)$ and let $u_1, u_2, v_1, v_2 \in \Sigma^*$, $c \in B$ be such that $u = u_1 a u_2$, $v = c v_1 b v_2$ (eventually $c = b$). Then we obtain by Prop. 12 with $(au_2, c) \in I$ that $a < c$, thus also $a < b = \max(B)$.

Conversely suppose that aub is a factor of z with $(au, b) \in I$, where z belongs to the right-hand side of Eq. (5). Then $a \leq \max(au) < \max(b) = b$ yields the claim.

Finally, using again Prop. 12 it is easy to check that $xyz \in \mathrm{LNF}$ and $(y, z) \in I$ imply $xz \in \mathrm{LNF}$.

Let us consider now the priority normal form used in [18]. Assume that the alphabet Σ is totally ordered by $<$. Then for $w \in \Sigma^+$ we denote by $\min(w)$ the lowest letter (w.r.t. $<$) occurring in w, i.e. $\min(w) = \min(\mathrm{alph}(w))$.

Definition 22. Let Σ be totally ordered by $<$. The *priority normal form* of $w \in \Sigma^*$, $\mathit{pnf}(w)$, is defined inductively as follows:

i) $\mathit{pnf}(1) = 1$.
ii) For $w \neq 1$ let $a = \min(w)$ and choose $v \equiv_I w$ such that $v = v_1 a v_2 a \ldots v_k a v_{k+1}$ with $|v_1 \cdots v_{k+1}|_a = 0$ and each $|v_i|$, $1 \leq i \leq k$, is minimal. That is, we choose a word v equivalent to w such that every occurrence of the minimal letter in w appears in v at the leftmost position. Note that this means that for every position in v_i, $1 \leq i \leq k$, there is a dependence path from this position to the ith occurrence of a (a dependence path is an increasing sequence of positions such that every two consecutive positions are labelled by a pair of dependent letters). Moreover, the lengths of v_i, $1 \leq i \leq k$, depend on w, only. We denote the vector $(|v_1|, \ldots, |v_k|)$ by $||w||$. The priority normal form of w, $\mathit{pnf}(w)$, is defined as

$$\mathit{pnf}(w) = \mathit{pnf}(v_1)\, a\, \mathit{pnf}(v_2)\, a \, \cdots \mathit{pnf}(v_k)\, a\, \mathit{pnf}(v_{k+1}) \, .$$

We denote the set of all words in priority normal form, $\{\mathit{pnf}(w) \mid w \in \Sigma^*\}$, by PNF.

We can characterize PNF by forbidden factors, similarly to Prop. 12:

Proposition 23. *Let Σ be totally ordered by $<$. Then a word $w \in \Sigma^*$ is in priority normal form if and only if for every factor uv of w with $(u,v) \in I$, $u \neq 1$ and $v \neq 1$, we have $\min(u) < \min(v)$.*

Proof. Suppose first that $w \in \mathrm{PNF}$, $w \neq 1$, and consider the decomposition $w = w_1 a w_2 a \ldots w_k a w_{k+1}$, with $|w_1 \cdots w_{k+1}|_a = 0$, where $a = \min(w)$. If $w \in a^+$ then the claim is immediately satisfied. So suppose that $|\mathrm{alph}(w)| > 1$ and consider a factor uv of w with $(u,v) \in I$, $u \neq 1$, $v \neq 1$. Let $b = \min(u)$, $c = \min(v)$ and assume that $u = u_1 b u_2$, $v = v_1 c v_2$ for some u_1, u_2, v_1, v_2. If bu_2v_1c is a factor of some w_i, then we have $b < c$ by induction on the size of $\mathrm{alph}(w)$. Otherwise suppose that b occurs in w_i and c occurs in w_j with $i < j$, and assume by contradiction that $b > c$. Then from the definition of b, c it follows that $a \in \{b, c\}$, thus $c = a$. Thus there is a dependence path from b to c, which contradicts the assumption $(u,v) \in I$.

For the converse consider $w = w_1 a w_2 a \dots w_k a w_{k+1} \neq 1$, with $|w_1 \cdots w_{k+1}|_a = 0$ and $a = \min(w)$, such that w satisfies the condition $\min(u) < \min(v)$ for all factors uv with non-empty u, v. For $w \in a^+$ it follows immediately that $w \in \mathrm{PNF}$. If $\mathrm{alph}(w) > 1$ then it suffices to show that $(|w_1|, \dots, |w_k|) = ||w||$ and then apply induction over the alphabet to each of the words w_i. Let us assume by contradiction that $|w_i|$ is not minimal, for some $1 \leq i \leq k$. Then let $b \in \Sigma$, $x, y \in \Sigma^*$ be such that $w_i a = xbya$ and there is no dependence path from b to a. Moreover, assume that $|y|$ is minimal with this property. Hence, $(b, ya) \in I$ and we may conclude by noting that $\min(b) = b > a = \min(ya)$ yields a contradiction to the assumption on w.

Proposition 24. *Let Σ be totally ordered by $<$ and let (Σ, I) be an independence alphabet. Then* PNF *is a set of left-factor normal forms for $\mathbb{M}(\Sigma, I)$.*

Proof. Let us consider the monoid $S = (\Sigma \dot\cup \{1\}, \cdot, 1)$ with $a \cdot b = \min(a,b)$. It is easy to check that together with $a \leq 1$ for all $a \in \Sigma$ we obtain an ordered monoid, $S = (S, <)$. Moreover, $h(a) = a$ yields $h(w) = \min(w)$ for every $w \in \Sigma^+$. Therefore, the characterization given by Prop. 23 yields directly condition (i) of Def. 19.

Consider now $xyz \in \mathrm{PNF}$ with $(x,y) \in I$. Assume that uv is a factor of xz, where $(u,v) \in I$, $u \neq 1$, $v \neq 1$. If uv is a factor of either x or z, then we directly conclude that $\min(u) < \min(v)$. Otherwise we distinguish two cases. If $v = v_1 v_2$ is such that uv_1 is a suffix of x and v_2 is a prefix of z, then $(u, v_1 y v_2) \in I$ and thus $\min(u) < \min(v_1 y v_2) \leq \min(v)$. Otherwise, suppose that $u = u_1 u_2$ is such that u_1 is a suffix of x and $u_2 v$ is a prefix of z, $u_2 \neq 1$. Then $(u_2, v) \in I$ yields $\min(u) \leq \min(u_2) < \min(v)$. By Prop. 23 we conclude that $xz \in \mathrm{PNF}$.

4.3 Normal Forms and Factorizations

The proposition below is a reformulation of Lem. 14 for left-factor normal forms. The proof is almost identical to the proof of Lem. 14, so we only point out the main observations needed for the result. We identify below the finite monoid S with the ordered set $\{1, 2, \dots, |S|\}$. Recall that $n = |\Sigma|$.

Proposition 25. *Let $\mathcal{N} \subseteq \Sigma^*$ be a set of left-factor normal form for $\mathbb{M}(\Sigma, I)$ via the finite ordered monoid $(S, <)$, $S = (S, \cdot, 1)$ with $s \leq 1$ for all $s \in S$, and the homomorphism $h : \Sigma^* \to S$. Let $u, v, w \in \mathcal{N}$ and $h = h(v)$. Then $uv \equiv_I w$ implies that some $1 \leq m \leq \frac{(n-1)(|S|-h)}{2} + 1$ and some words $u_1, \dots, u_m, v_1, \dots, v_m \in \mathcal{N}$ exist with*

$$\begin{array}{l} u \equiv_I u_1 \cdots u_m \quad (\textit{trace equality}), \\ v = v_1 \cdots v_m \quad (\textit{word equality}), \\ w = u_1 v_1 \cdots u_m v_m \quad (\textit{word equality}), \\ u_i \neq 1, \ \textit{for all} \ 1 < i \leq m, \\ v_j \neq 1 \ \textit{for all} \ 1 \leq j < m, \\ v_j \in I(u_{j+1} \cdots u_m) \ \textit{for all} \ 1 \leq j < m, \\ h(u_i) > h \ \textit{for all} \ 1 < i \leq m. \end{array}$$

Proof. Let $u_i, v_j \in \Sigma^*$ be defined as in the proof of Lem. 14. We first show that we have $h(u_{j+1}) > h$ for all $1 \leq j < m$. By Def. 19 $(v_j, u_{j+1}) \in I$ and $w \in \mathcal{N}$ imply $h(u_{j+1}) > h(v_j)$. Moreover, using that $s \leq 1$ for all s we obtain

$$h(u_{j+1}) > h(v_j) \geq h(v).$$

In a further step we show that $v_1 \cdots v_m \in \mathcal{N}$, which directly implies the word equality $v = v_1 \cdots v_m$. Since $w \in \mathcal{N}$ and $v_j \in I(u_{j+1} \cdots u_m)$ we may apply condition (ii) of Def. 19 for $u_1, \dots u_m$ and we obtain the claimed result.

Finally, it remains to show that m is bounded. To see this, we assume that $m > \frac{(n-1)(|S|-h)}{2} + 1$ and we proceed as in the proof of Lem. 14 with some positions $1 \leq i < j \leq m$ satisfying $j - i \geq \frac{|S|-h}{2}$ and $I(u_{i+1} \cdots u_m) = I(v_{j+1} \cdots v_m)$. Similarly, we obtain

$$h(v_i) < h(u_{i+1}) < h(v_{i+1}) < \cdots < h(v_j) < h(u_{j+1}).$$

This chain of inequalities yields together with $s \leq 1$ for all s:

$$h = h(v) \leq h(v_i) \leq (|S| - 2(j - i + 1) + 1),$$

hence $j - i \leq (|S| - h + 1)/2 - 1$, contradiction.

5 Trace Equations and Normal Forms

In this section we conclude the proof of Thm. 7 and we consider some immediate extensions of the main result. For the rest of the sequel we assume for simplicity that $\mathcal{N} = \mathrm{LNF}$. We could work instead with the priority normal form from Sect. 4.2 or with any set of left- or right-factor normal form.

Proof of Thm. 7. We can apply Prop. 16 to the trace equality $uv \equiv_I w$ and we obtain for some $\alpha \leq n^n n!$:

$$\begin{array}{l} u = u_1 \cdots u_\alpha, \\ v = v_{j_1} \cdots v_{j_\alpha}, \\ w = u_1 v_1 \cdots u_\alpha v_\alpha, \end{array}$$

for some permutation $(j_1, \dots, j_\alpha)$ of $(1, \dots, \alpha)$ and words $u_i, v_j \in \mathrm{LNF}$ satisfying

$$\begin{array}{l} (u_i, v_j) \in I \text{ for all } i > j \text{ and} \\ (v_{j_p}, v_{j_q}) \in I \text{ for all } p, q \text{ with } (j_p - j_q)(p - q) < 0. \end{array}$$

Of course, we can even assume that we have $\alpha = n^n n!$ above, by adding some empty factors u_i, v_j. The set $\mathcal{W}_\alpha(x,y,z)$ of systems of word equations with commutation constraints is then directly defined as consisting of all systems S of the form

$$\begin{array}{l} x = x_1 \cdots x_\alpha \ \& \\ y = y_{j_1} \cdots y_{j_\alpha} \ \& \\ z = x_1 y_1 \cdots x_\alpha y_\alpha \ \& \\ (x_i, y_j) \in I \text{ for all } i > j \ \& \\ (y_{j_p}, y_{j_q}) \in I \text{ for all } p, q \text{ with } (j_p - j_q)(p - q) < 0\,, \end{array}$$

for all permutations $(j_1, \ldots, j_\alpha)$ of $(1, \ldots, \alpha)$. The mapping σ with $\sigma(x) = u$, $\sigma(y) = v$, $\sigma(z) = w$, $\sigma(x_i) = u_i$, and $\sigma(y_j) = v_j$ for $1 \le i, j \le \alpha$, satisfies the requirements of the theorem. This concludes the proof.

We consider in the following two easy generalizations of Cor. 10. First, recall the generalization of Makanin's theorem to word equations with regular constraints given by Schulz [22]. However, in the case of trace equations we cannot add arbitrary regular constraints while preserving decidability. This is due to a result of Aalbersberg and Hoogeboom stating that it is in general undecidable whether for given regular languages $L, K \subseteq \Sigma^*$ some words $u \in L$, $v \in K$ exist such that $u \equiv_I v$ [1]. Moreover, this question can be rephrased in terms of trace equations, e.g. as $x \equiv y$ with the constraints $L_x = L$, $L_y = K$. However, we preserve decidability if we restrict to a subclass of regular constraints, to I-closed regular constraints:

Definition 26. Let $L \subseteq \Sigma^*$ be a language and let I be an independence relation. Then L is called I-closed if for each $u \in L$ and $v \equiv_I u$, we have also $v \in L$.

Definition 27. A system T of trace equations over (Σ, I), Ω with I-closed regular constraints consists of a system T_0 of trace equations, together with I-closed regular languages L_x for each $x \in \Omega$.

A mapping $\sigma : \Omega \to \Sigma^*$ is called a solution for T if

i) σ is a solution for T_0, and
ii) $\sigma(x) \in L_x$ is satisfied for all $x \in \Omega$.

Adding I-closed constraints $L_x \subseteq \Sigma^*$, $x \in \Omega$, to a system of trace equations T means that we preserve the property that for each solution σ for $T' = (T, (L_x)_{x \in \Omega})$ we also have $\nu_{\mathcal{N}} \circ \sigma$ as a solution for T', for e.g. $\mathcal{N} = \mathrm{LNF}$. Therefore, Cor. 9 can be immediately extended to systems with I-closed regular constraints:

Corollary 28. *Let T be a system of trace equations with unknowns Ω over (Σ, I), with I-closed regular constraints. Then a set of unknowns $\Omega' \supseteq \Omega$ and a finite set $\mathcal{W}$ of systems of word equations with unknowns Ω' and regular constraints over Σ can be effectively constructed such that we have:*

i) For every solution σ for T the mapping $\nu_{\mathcal{N}} \circ \sigma$ can be extended to a solution for some $S \in \mathcal{W}$.
ii) For every $S \in \mathcal{W}$ and every solution σ for S, $\sigma|_\Omega$ is a solution for T.

Our second generalization concerns the possibility of using all Boolean connectives over trace equations. Since the class of I-closed regular languages over Σ^* forms a Boolean algebra (as the family of recognizable subsets of $\mathbb{M}(\Sigma, I)$, see e.g. [9]) we need only to consider negations of trace equations.

Lemma 29. *Let $L \equiv R$ be a trace equation with unknowns Ω over (Σ, I). There is a natural bijection between the solutions of the inequality*

$$\mathbf{not}(L \equiv R)$$

and all solutions of all systems of trace equations with constraints of the form

$$L \equiv xy \ \ \& \ \ R \equiv xz \ \ \& \ \ \mathit{init}(y) = A \ \ \& \ \ \mathit{init}(z) = B\,, \tag{6}$$

where x, y, z denote new unknowns and we consider in (6) all alphabets $A, B \subseteq \Sigma$ such that $A \cap B = \emptyset$ and $A \cup B \neq \emptyset$.

Before giving the proof, let us explain how constraints of the form $\mathrm{init}(y) = A$ can be expressed by trace equations with commutation constraints. First, an equation of the form $y \equiv ay'$ ensures that $a \in \mathrm{init}(y)$. Second, the following positive formula over trace equations and commutation constraints ensures that $a \notin \mathrm{init}(y)$:

$$(y, a) \in I \ \vee \bigvee_{b \neq a,\, (a,b) \notin I} (y \equiv y'by'' \ \&\ (y', a) \in I)\,,$$

where y', y'' denote new variables (recall that I is irreflexive).

Proof. Suppose that σ is such that $\sigma(L) \not\equiv_I \sigma(R)$. Let s denote the longest word such that for some $t, u \in \Sigma^*$:

$$\sigma(L) \equiv_I st \text{ and } \sigma(R) \equiv_I su\,.$$

Due to the maximality of s we obtain $\mathrm{init}(t) \cap \mathrm{init}(u) = \emptyset$. Moreover, $tu \neq 1$, since $st \not\equiv_I su$. Hence, $\sigma' = \sigma \cup \{x \mapsto s, y \mapsto t, z \mapsto u\}$ is a solution for one of the new subformulas (6). Conversely, for every solution σ' of (6) one has $\sigma'(x)\sigma'(y) \not\equiv_I \sigma'(x)\sigma'(z)$, hence $\sigma'(L) \not\equiv_I \sigma'(R)$.

The previous lemma shows that the existential first-order theory of concatenation modulo partial commutation is decidable. More precisely, we can consider a first-order formulas F over atomic predicates which are either trace equations with unknowns Ω over (Σ, I) or regular constraints $x \in L_x$ for $x \in \Omega$, where L_x is I-closed. An assignment $\sigma : \Omega \to \Sigma^*$ is called a solution for F if F evaluates to true when replacing each trace equation $L \equiv R$ by the truth value of $\sigma(L) \equiv_I \sigma(R)$, resp. each constraint $x \in L_x$ by the truth value of $\sigma(x) \in L_x$.

Corollary 30. *Let (Σ, I) be an independence alphabet. It is decidable whether a first-order formula over atomic predicates which are either trace equations over (Σ, I) or regular I-closed constraints, is satisfiable.*

We conclude this section by showing that commutation constraints $(x, y) \in I$ can be expressed by word equations with regular constraints over the same alphabet Σ, i.e. without introducing a new letter as in Sect. 3. We assume without restriction that (Σ, I) contains no letter a with $(a, b) \in I$ for all $b \neq a$. Otherwise, we can reduce the solvability of trace equations over (Σ, I) to the solvability of trace equations over (Σ', I'), $\Sigma' = \Sigma \setminus \{a\}$, $I' = I \cap (\Sigma' \times \Sigma')$, and the solvability of a system of linear equations over $\mathbb{N}$. More generally, the solvability of any system of trace equations over the direct product $\mathbb{M}(\Sigma, I) \times \mathbb{M}(\Sigma', I')$ can be reduced to the solvability of two systems, one over $\mathbb{M}(\Sigma, I)$, the other one over $\mathbb{M}(\Sigma', I')$.

Let $\Sigma = \{a_1, \ldots, a_n\}$ and define $\alpha, \beta \in \Sigma^*$ as

$$\alpha = a_1 \cdots a_n \qquad \beta = a_n \cdots a_1 \,.$$

Consider the regular languages $L_i = \bigcup_{A,B \subseteq \Sigma, A \times B \subseteq I} A^* \alpha^i \beta^i B^*$, for $i \geq 1$.

Lemma 31. *Let $u, v \in \Sigma^*$. Then $(u, v) \in I$ is equivalent to the conjunction of three three conditions $u\alpha^i\beta^i v \in L_i$, for $i = 1, 2, 3$.*

Proof. Let $u\alpha^i\beta^i v \in L_i$ for $i = 1, 2, 3$. We denote below by $\pi_{a,b} : \Sigma^* \to \{a, b\}^*$ the projection morphism to the alphabet $\{a, b\} \subseteq \Sigma$. Moreover, let $u' = \pi_{a,b}(u)$, $v' = \pi_{a,b}(v)$.

By contradiction we may suppose that $a \neq b$, $(a, b) \notin I$ exist such that $u' \neq 1$ and $v' \neq 1$ (we use the assumption on (Σ, I)). We have $u'(ab)^i(ba)^i v' \in \pi_{a,b}(L_i)$. For $i > 0$ we obtain that either $(ab)^i b$ is a prefix of u', or $b(ba)^i$ is a suffix of v'. This is seen by noting that each word in $\pi_{a,b}(L_i)$ has $(ab)^i(ba)^i$ either as a prefix or as a suffix, and by considering the possible self-overlaps of $(ab)^i(ba)^i$. But every combination of the two cases above for $i = 1, 2, 3$ yields a contradiction, hence $u' = 1$ or $v' = 1$.

Proposition 32. *Let Ω be a set of unknowns, $x, y \in \Omega$. There is a natural bijection between the solutions for a commutation constraint $(x, y) \in I$ and the solutions for the following system of word equations with regular constraints over Σ:*

$$\bigwedge_{i=1}^{3} (z_i = x(ab)^i(ba)^i y \quad \& \quad z_i \in L_i) \,,$$

where $z_1, z_2, z_3 \notin \Omega$ denote new unknowns.

6 Mixed Equation Systems

Knowing that we can decide the solvability of systems of trace equations, resp. of word equations, it is natural to ask whether the decidability is preserved when we consider both equation types in one system. Actually, our reduction from trace equations to word equations could be also rephrased in such a way that we have to consider mixed systems of word and trace equations, as it was e.g. done in [8]. Fortunately, the mixed systems arising in that reduction process have the property that for each solution of the system, the normal form of the solution is still a solution. This property is crucial for the decidability result. We show below that this property cannot be verified in general. Moreover, we show that for an arbitrary system of trace and word equations it is undecidable whether the system is solvable. A similar proof idea was used in [4,11] where it is shown that adding two length predicates to word equations (e.g. adding requirements of the form $|x|_a = |y|_a$, $|x|_b = |y|_b$ for $a \neq b$) makes the question of solvability undecidable.

Proposition 33. *Let $\mathbb{M}(\Sigma, I)$ be a trace monoid and let S be a system of trace equations and word equations. Then:*

i) It is undecidable whether S has a solution.
ii) Let $\mathcal{N} \subseteq \Sigma^*$ be a set of left-factor (or right-factor) normal forms for $\mathbb{M}(\Sigma, I)$ and let $\nu : \Sigma^* \to \mathcal{N}$ denote the canonical mapping. It is undecidable whether for every solution $\sigma : \Omega \to \Sigma^*$ for S the mapping $\nu \circ \sigma$ is also a solution.

Proof. For the first assertion we reduce Hilbert's Tenth Problem which is known to be undecidable [19,16] to the solvability of a system of word equations and trace equations over (Σ, I) with $\Sigma = \{a, b\}$ and $I = \{(a, b), (b, a)\}$. Since every Diophantine equation E can be translated into a system with equations of the form $z = x + y$, $z = x \cdot y$, $x = y$, resp. $x = 1$ it suffices to transform each of these simple equations into conjunctions of word and trace equations. First we associate to each variable x occurring in E three variables $x_a, x_b, x_{a,b} \in \Omega$ together with two word equations and one trace equation:

$$x_a a = a x_a \quad \& \quad x_{a,b} ab = ab x_{a,b} \quad \& \quad x_{a,b} \equiv x_a x_b \,.$$

Note that every solution $\sigma : \Omega \to \Sigma^*$of the above system is of the form $\sigma(x_a) = a^i$, $\sigma(x_b) = b^i$, $\sigma(x_{a,b}) = (ab)^i$ for some $i \geq 0$. That is, $|x_a| = |x_b|$ will encode the value of the variable x of E.

Equations of the form $z = x + y$ will be encoded straightforwardly into a word equation $z_a = x_a y_a$ (analogously for $x = y$, $x = 1$). An equation of the form $z = x \cdot y$ will be encoded by two equations:

$$z_{a,b} x_a b = x_a b z_{a,b} \quad \& \quad z_{a,b} \equiv z_a y_b$$

Note that $a^* b$ consists only of primitive words (that is, words which cannot be written as r^n, for some word r and $n > 1$). Therefore every solution to the first equation above yields $z_{a,b} \in (x_a b)^*$. Thus, the system $S(E)$ of trace equations and word equations defined above is solvable if and only if the Diophantine equation E has a solution.

For the second assertion of the proposition note that given a Diophantine equation E with constant coefficient non-zero, E has a solution if and only if E

has a non-singular solution σ, i.e. a solution σ in nonnegative integers such that $\sigma(x) \neq 0$ for some unknown x. Moreover, the equivalent system of equations of the form $z = x + y$, $z = x \cdot y$, $x = y$, $x = 1$ also satisfies the property of having either no solution or a non-singular solution. Assume that $ab \in \mathcal{N}$ (otherwise we interchange a, b in the mixed system of equations given above). For each unknown $x_{a,b}$ we add an equation $x'_{a,b} = abx_{a,b}$, where $x'_{a,b}$ is a new unknown. Due to the unknowns $x'_{a,b}$ we have that $S(E)$ has either no solution or it has a solution σ with $\sigma(x'_{a,b}) \in (ab)^2(ab)^*$ for some $x'_{a,b}$. Moreover, every solution σ satisfies $\sigma(x'_{a,b}) \in (ab)^*$, for every x. However, by the definition of right-factor resp. left-factor normal forms we have for each word $w \in \{a,b\}^*$ either $\nu(w) \in a^*b^*$ or $\nu(w) \in b^*a^*$. Therefore, $S(E)$ satisfies the property that for every solution σ, $\nu \circ \sigma$ is also a solution if and only if E has no solution.

Remark 34. Prop. 33 also shows that there exists no finite set $\mathcal{W}$ of systems of word equations over $\Sigma = \{a,b\}$ and unknowns $\Omega \supseteq \{x,y,z\}$ such that the solutions of $\mathcal{W}$ are exactly the mappings $\sigma : \Omega \to \{a,b\}^*$ satisfying

$$\sigma(x) = a^i,\ \sigma(y) = b^j,\ \sigma(z) = w \text{ with } |w|_a = i, |w|_b = j\,,$$

for $i, j > 0$. Otherwise we could replace the trace equations $x_{a,b} \equiv x_a x_b$ in the proof of Prop. 33 by formulas over word equations, thereby obtaining the decidability of Hilbert's Tenth Problem.

7 Parallel Complexity of Computing Normal Forms

Computing the lexicographical normal form of a given word is easy, it can be performed deterministically in linear time. In this section we are concerned with the parallel complexity of computing lexicographical normal forms. We consider uniform circuit complexity classes like AC^0 and TC^0. Let $f : \Sigma^* \to \Sigma^*$ be a function such that $|f(w)| \leq p(m)$ for all w with $|w| = m$, where p is a polynomial. Let $k \geq 0$. Then f is AC^k-computable if there is a family $(C_m)_{m\geq 0}$ of polynomial-size circuits of depth $O(\log^k(m))$ with AND and OR gates of unbounded fan-in/out and unary NOT gates, such that $C_{|w|}$ computes $f(w)$ for all $w \in \Sigma^*$. A function f is TC^k-computable if there is a family of circuits as above which in addition to AND, OR and NOT gates contain MAJORITY gates of unbounded fan-in/out. A MAJORITY gate yields 1 if and only if more than half of its inputs are 1. In order to be able to deal with arbitrary alphabets Σ one usually assumes that the circuits have special input/output gates testing $x = a$ for each input position x and letter $a \in \Sigma$ (analogously for the outputs). Uniformity means that given $n \geq 0$ (a fixed coding of) the circuit C_n can be easily computed (e.g. in logarithmic space). It is well-known that $\mathrm{AC}^k \subseteq \mathrm{TC}^k \subseteq \mathrm{AC}^{k+1}$, $k \geq 0$. For more details about circuit complexity see e.g. [23]. With Prop. 16 we obtain

Corollary 35. *Let (Σ, I) denote a fixed independence alphabet. Then we can compute $\ell\mathrm{nf}(uv)$ on input $u, v \in \mathrm{LNF}$ in uniform AC^0.*

Proof. As it suffices to take $\alpha = n^n n!$ in Prop. 16 we build a subcircuit for each permutation $(j_1, \ldots, j_\alpha)$ of $(1, \ldots, \alpha)$ and each pair of factorizations $u = u_1 \cdots u_\alpha$, $v = v_{j_1} \cdots v_{j_\alpha}$. The subcircuit checks that $(u_i, v_j) \in I$ for $i > j$ resp. that $(v_{j_p}, v_{j_q}) \in I$ is satisfied for all p, q with $(p-q)(j_p - j_q) < 0$. Moreover, it checks that $u_1 v_1 \cdots u_\alpha v_\alpha$ is in lexicographical normal form. The last test is in AC^0 by Prop. 12. If the test is positive, this is also the output of the whole circuit. The size of the circuit is a polynomial in $|uv|$, since α is a constant.

We could apply Cor. 35 in order to compute the function ℓnf by divide-and-conquer in AC^1. However, we can compute ℓnf in $\mathrm{TC}^0 \subseteq \mathrm{AC}^1$ as shown below. This result can be compared with the fact that the equivalence $s \equiv_I t$ of two words $s, t \in \Sigma^*$ can be verified in uniform TC^0, too [2].

Proposition 36. *Let (Σ, I) denote a fixed independence alphabet. Then the mapping $\ell\mathrm{nf} : \Sigma^* \to \mathrm{LNF}$ is computable in uniform TC^0.*

Proof. For $w \in \Sigma^+$ and $1 \leq i, j \leq |w|$ let $\mathrm{lex}(i, j)$ denote a predicate which is satisfied by w if and only if the ith position of w comes before the jth position of w in the lexicographical normal form of w, $\ell nf(w)$. By [12, Prop. 3] there exists a first-order definable predicate $\mathrm{lex}(i, j)$ as above. But first-order definable languages belong to AC^0, i.e. their characteristic function is AC^0-computable (see e.g. [23]). Finally, the ith position of a word w occurs as the jth position of $\ell nf(w)$ if and only if $j = \#\{k \mid \mathrm{lex}(k, i)\} + 1$. The last statement can be verified in TC^0 by using a MAJORITY gate for counting the number of values of k satisfying $\mathrm{lex}(k, i)$.

Remark 37. Note that ℓnf cannot be computed in AC^0. This is seen by noting that for the independence alphabet $\{a, b\}$ with $(a, b) \in I$, computing $\ell nf(w)$ for some $w \in \{a, b\}^*$ is equivalent to counting, e.g. $|w|_a$. But this is a problem known not to belong to AC^0.

References

1. IJ. J. Aalbersberg and H. J. Hoogeboom. Characterizations of the decidability of some problems for regular trace languages. *Mathematical Systems Theory*, 22:1–19, 1989.
2. C. Àlvarez and J. Gabarró. The parallel complexity of two problems on concurrency. *Information Processing Letters*, 38:61–70, 1991.
3. A. V. Anisimov and D. E. Knuth. Inhomogeneous sorting. *International Journal of Computer and Information Sciences*, 8:255–260, 1979.
4. J. R. Büchi and S. Senger. Definability in the existential theory of concatenation and undecidable extensions of this theory. *Zeitschr. f. math. Logik und Grundlagen d. Math.*, 34:337–342, 1988.
5. P. Cartier and D. Foata. *Problèmes combinatoires de commutation et réarrangements.* Number 85 in Lecture Notes in Mathematics. Springer, Berlin-Heidelberg-New York, 1969.

6. C. Choffrut. Combinatorics in trace monoids I. In V. Diekert and G. Rozenberg, editors, *The Book of Traces*, chapter 3, pages 71–82. World Scientific, Singapore, 1995.
7. C. Choffrut and J. Karhumäki. Combinatorics of words. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, pages 329–438. Springer, Berlin-Heidelberg-New York, 1997.
8. V. Diekert, Yu. Matiyasevich, and A. Muscholl. Solving trace equations using lexicographical normal forms. In P. Degano et al., editors, *Proceedings of the 24th International Colloquium on Automata, Languages and Programming (ICALP'97), Bologna (Italy) 1997*, number 1256 in Lecture Notes in Computer Science, pages 336–347, Berlin-Heidelberg-New York, 1997. Springer.
9. V. Diekert and G. Rozenberg, editors. *The Book of Traces.* World Scientific, Singapore, 1995.
10. C. Duboc. On some equations in free partially commutative monoids. *Theoretical Computer Science*, 46:159–174, 1986.
11. V. G. Durnev. On equations in free semigroups and groups. *Matematicheskie Zametki*, 16(5):717–724, 1974. (In Russian) English translation in: Math. Notes of the Acad. of Sci. of the USSR 16 (1975) 1024–1028.
12. W. Ebinger and A. Muscholl. Logical definability on infinite traces. *Theoretical Computer Science*, 154:67–84, 1996.
13. J. Karhumäki, W. Plandowski, and F. Mignosi. The expressibility of languages and relations by word equations. In P. Degano et al., editors, *Proceedings of the 24th International Colloquium on Automata, Languages and Programming (ICALP'97), Bologna (Italy) 1997*, number 1256 in Lecture Notes in Computer Science, pages 98–109, Berlin-Heidelberg-New York, 1997. Springer.
14. R. M. Keller. Parallel program schemata and maximal parallelism I. Fundamental results. *Journal of the Association for Computing Machinery*, 20(3):514–537, 1973.
15. G. S. Makanin. The problem of solvability of equations in a free semigroup. *Math. Sbornik*, 103:147–236, 1977. English transl. in Math. USSR Sbornik 32 (1977).
16. Yu. Matiyasevich. *Hilbert's tenth problem.* Foundations of computing. MIT Press, Cambridge, Mass., 1993.
17. Yu. Matiyasevich. Reduction of trace equations to word equations, 1996. Talk given at the "Colloquium on Computability, Complexity, and Logic", December 5-6, 1996, Institut für Informatik, Universität Stuttgart, Germany.
18. Yu. Matiyasevich. Some decision problems for traces. In S. Adian and A. Nerode, editors, *Proceedings of the 4th International Symposium on Logical Foundations of Computer Science (LFCS'97), Yaroslavl, Russia, July 6–12, 1997*, number 1234 in Lecture Notes in Computer Science, pages 248–257, Berlin-Heidelberg-New York, 1997. Springer.
19. Yu. V. Matiyasevich. Enumerable sets are Diophantine. Doklady Akademii Nauk SSSR, 191(2):279–282, 1970. English transl. in Soviet Mathematics. Doklady, vol. 11, no. 2, 1970.
20. A. Mazurkiewicz. Concurrent program schemes and their interpretations. DAIMI Rep. PB 78, Aarhus University, Aarhus, 1977.
21. J.-E. Pin. Syntactic semigroups. In G. Rozenberg and A. Salomaa, editors, *Handbook of Formal Languages*, volume 1, pages 679–738. Springer, Berlin-Heidelberg-New York, 1997.
22. K. U. Schulz. Makanin's algorithm for word equations — Two improvements and a generalization. In K. U. Schulz, editor, *Word Equations and Related Topics*, number 572 in Lecture Notes in Computer Science, pages 85–150, Berlin-Heidelberg-New York, 1991. Springer.

23. H. Straubing. *Finite automata, formal logic, and circuit complexity*. Birkhäuser, 1994.