On “simple” undecidable fragments of positive and elementary theories of free semigroups of finite or countable rank

V G Durnev, O V Zetkina and A I Zetkina
P. G. Demidov Yaroslavl’ University, 150003 Yaroslavl’, Russia
E-mail: ovzetkina@yandex.ru

Abstract. We establish algorithmic undecidability of “simple” (with respect to the quantifier-prefix and the quantifier-free part of the formulas) fragments of positive theories of finitely generated semigroups and elementary theory of a free countably generated semigroup.

1. Introduction

Although the word problem and its decidability has been ever-present in mathematics [1], the combinatorics of words has wide recent applications in number theory, coding, computer science and technologies (structure matching and searching, data structure indexing), linguistics (statistical and symbolic language processing) and bioinformatics (probabilistic models for biological sequences such as DNA) [2]. Word equations are important in string solvers which are essential component in automated systems for information extraction (document spanners) [3], as well as in many formal verification, security analysis, and bug-finding tools [4]. Such applications require deep understanding of the boundary for decidability for fragments of the theory of word equations [5] as well as evaluating the complexity of the corresponding algorithms [6, 7].

Let us denote by $S_m$ the free semigroup of rank $m$ with free generators $a_1, \ldots, a_m$. For the case $m = 2$ we write $a, b$ instead of $a_1, a_2$, and for the case $m = 3$ we write $a, b$ and $c$ instead of $a_1, a_2$ and $a_3$ correspondingly. Note that $S_1$ is a cyclic semigroup. Henceforward we deal with noncyclic (noncommutative) semigroups $S_m$ only, i. e. we will assume $m \geq 2$. For some reasons, of special interest is the “boundary” case – free semigroup $S_2 = \langle a, b \rangle$ with two free generators $a$ and $b$.

Elementary theory of free semigroup $S_m$ is the set of all closed (i. e. those not containing free occurrences of variables) formulas $\Phi$ of the type

$$(Q_1 x_1)(Q_2 x_2)\ldots(Q_n x_n) \Psi,$$

where $\Psi = \bigwedge_{i=1}^k \left( \bigwedge_{j \in A_i} w_{ij} = u_{ij} \right) \land \left( \bigwedge_{t \in B_i} v_{it} \neq z_{it} \right)$,

$w_{ij}, u_{ij}, v_{it}, z_{it}$ are words in the alphabet $\{ x_1, \ldots, x_n, a_1, \ldots, a_m \}$,

$A_i$ and $B_i$ are sets (possibly empty), and $Q_1, \ldots, Q_n$ are the quantifiers $\forall$ or $\exists$.
which are true on the free semigroup $S_m$.

We call $(Q_1 x_1)(Q_2 x_2)\ldots(Q_n x_n)$ the quantifier prefix of the formula $\Phi$, and $Q_1 Q_2 \ldots Q_n$ – the type of quantifier prefix, and $\Psi$ – quantifier-free part of the formula $\Phi$.

A formula $\Phi$ is called positive if its quantifier-free part, $\Psi$, does not contain negation sign, i.e. if it has the form:

$$\bigvee_{i=1}^k \left( \bigwedge_{j \in A_i} w_{ij} = u_{ij} \right).$$

The positive theory of a free semigroup $S_m$ is the set of all closed positive formulas $\Phi$ which are true on the free semigroup $S_m$.

The study of elementary theory of free non-commutative semigroup began with the work by W. Quine [8] where he proved algorithmic undecidability of the elementary theory of a non-cyclic free semigroup.

Because of the equivalence

$$U \neq V \iff (\exists x)(\exists y)(\exists z) \left( \bigvee_{i=1}^m (U = V a_i x \lor V = U a_i x) \lor \left( \bigvee_{i,j=1, i \neq j}^m U = x a_i y \land V = x a_j z \right) \right),$$

which holds for any two elements $U$ and $V$ of a semigroup $S_m$, algorithmic undecidability of the positive theory of a non-cyclic free semigroup follows directly from Quine’s result, although this fact was not mentioned in Ref. [8].

This result was strengthened substantially in Ref. [9] where the following fact was proved:

It is possible to construct such one-parameter family of formulas

$$\exists y)(\forall z)(\exists x_1)(\exists x_2)(\exists x_3) \left( \bigvee_{i=1}^{14} w_i(x, y, z, x_1, x_2, x_3, a, b) = u_i(x, y, z, x_1, x_2, x_3, a, b) \right),$$

with single parameter $x$, so that there would not exist a deciding algorithm to determine, for a given arbitrary word $A$ (element of free semigroup $S_2$), whether the positive formula

$$\exists y)(\forall z)(\exists x_1)(\exists x_2)(\exists x_3) \left( \bigvee_{i=1}^{14} w_i(A, y, z, x_1, x_2, x_3, a, b) = u_i(A, y, z, x_1, x_2, x_3, a, b) \right).$$

is true on this free semigroup $S_2$.

It was natural to pose the question: is it possible to simplify the quantifier-free part of the above formula? N. K. Kossovsky [10] constructed the formula $DK(x, y, z, v)$ of the type

$$\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4) w(x, y, z, v, x_1, x_2, x_3, x_4, a, b) = u(x, y, z, v, x_1, x_2, x_3, x_4, a, b)$$

such that the following equivalence holds for arbitrary elements $A$, $B$, $C$ and $D$ of free semigroup $S_m$:

$$A = B \lor C = D \iff S_m \models (\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4) w(A, B, C, D, x_1, x_2, x_3, x_4, a, b) = u(A, B, C, D, x_1, x_2, x_3, x_4, a, b).$$

This equivalence allows one to exclude the disjunction sign $\lor$ from the quantifier-free part of the considered formulas, however, the quantifier prefix being substantially complicated.
2. Setting of the problem and Methods
The above result by N. K. Kossovsky was somewhat strengthened in Ref. [11] where the formula
\[ DD(x, y, z, v) \] of the type

\[
(\exists x_1)(\exists x_2) w(x, y, z, v, x_1, x_2, a, b) = u(x, y, z, v, x_1, x_2, a, b)
\]

was constructed so that, for arbitrary elements \( A, B, C \) and \( D \) of free semigroup \( S_m \), the following equivalence holds:

\[
A = B \lor C = D \iff S_m \models (\exists x_1)(\exists x_2) w(A, B, C, D, x_1, x_2, a, b) = u(A, B, C, D, x_1, x_2, a, b).
\]

This allows one to simplify the quantifier-free part of the considered formulas, however, the quantifier prefix being obviously complicated. The following assertion holds:

*It is possible to construct such one-parameter family of formulas*

\[
(\exists y)(\forall z)(\exists x_1)(\exists x_2) \ldots (\exists x_{11}) w(x, y, z, x_1, x_2, \ldots, x_{11}, a, b) = u(x, y, z, x_1, x_2, \ldots, x_{11}, a, b),
\]

with single parameter \( x \), so that there would not exist a deciding algorithm to determine, for a given arbitrary word \( A \) (element of free semigroup \( S_2 \)), whether the positive formula

\[
(\exists y)(\forall z)(\exists x_1)(\exists x_2) \ldots (\exists x_{11}) w(A, y, z, x_1, x_2, \ldots, x_{11}, a, b) = u(A, y, z, x_1, x_2, \ldots, x_{11}, a, b).
\]

is true on this free semigroup \( S_2 \).

In Ref. [12], the formula \( DKMP_n(x, y, z, v) \) of the type

\[
(\exists x_1)(\exists x_2) w(y_1, x_1, x_2, a, b) = u(y_1, z_1, \ldots, y_n, z_n, x_1, x_2, a, b)
\]

was constructed such that for arbitrary elements \( A_1, B_1, \ldots, A_n \) and \( B_n \) of free semigroup \( S_m \) the following equivalence holds:

\[
A_1 = B_1 \lor A_2 = B_2 \lor \ldots \lor A_n = B_n \iff S_m \models (\exists x_1)(\exists x_2) w(A_1, B_1, A_2, B_2, \ldots, A_n, B_n, x_1, x_2, a, b) = u(A_1, B_1, A_2, B_2, \ldots, A_n, B_n, x_1, x_2, a, b).
\]

This fact allows one to simplify the quantifier-free part of the considered formulas at the cost of insignificant complication of the quantifier prefix. Therefore we obtain the following theorem.

**Theorem 1.** It is possible to construct such one-parameter family of formulas

\[
(\exists y)(\forall z)(\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4)(\exists x_5) w(x, y, z, x_1, x_2, x_3, x_4, x_5, a, b) = u(x, y, z, x_1, x_2, x_3, x_4, x_5, a, b),
\]

with single parameter \( x \), so that there would not exist a deciding algorithm to determine, for a given arbitrary word \( A \) (element of free semigroup \( S_2 \)), whether the positive formula

\[
(\exists y)(\forall z)(\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4)(\exists x_5) w(A, y, z, x_1, x_2, x_3, x_4, x_5, a, b) = u(A, y, z, x_1, x_2, x_3, x_4, x_5, a, b).
\]

is true on this free semigroup \( S_2 \).
S. S. Marchenkov [13] constructed such one-parameter family of formulas of the type

\[(\forall y)(\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4) \left( \bigvee_{i=1}^{m} w_i(x, y, x_1, x_2, x_3, x_4, a, b) = u_i(x, y, x_1, x_2, x_3, x_4, a, b) \right),\]

that there would not exist a deciding algorithm to determine, for a given arbitrary natural \(k\), whether the positive formula

\[(\forall y)(\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4) \left( \bigvee_{i=1}^{m} w_i(a^k, y, x_1, x_2, x_3, x_4, a, b) = u_i(a^k, y, x_1, x_2, x_3, x_4, a, b) \right),\]

is true on this free semigroup \(S_2\).

The results of the Refs. [9] and [13] were further strengthened in Ref. [11] where a one-parameter family of formulas of the type

\[(\forall y)(\exists x_1)(\exists x_2)(\exists x_3) \left( \bigvee_{i=1}^{m} w_i(x, y, x_1, x_2, x_3, a, b) = u_i(x, y, x_1, x_2, x_3, a, b) \right),\]

was constructed so that there would not exist a deciding algorithm to determine, for a given arbitrary natural \(k\), whether the positive formula

\[(\forall y)(\exists x_1)(\exists x_2)(\exists x_3) \left( \bigvee_{i=1}^{m} w_i(a^k, y, x_1, x_2, x_3, a, b) = u_i(a^k, y, x_1, x_2, x_3, a, b) \right)

is true on this free semigroup \(S_2\).

Using the formula (1) from Ref. [12] to remove the disjunction from the quantifier-free part, we obtain the following theorem.

**Theorem 2.** It is possible to construct such one-parameter family of formulas

\[(\forall y)(\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4)(\exists x_5) w(x, y, x_1, x_2, x_3, x_4, x_5, a, b) = u(x, y, x_1, x_2, x_3, x_4, x_5, a, b),\]

with single parameter \(x\), so that there would not exist a deciding algorithm to determine, for a given arbitrary naturel \(k\), whether the positive formula

\[(\forall y)(\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4)(\exists x_5) w(a^k, y, x_1, x_2, x_3, x_4, x_5, a, b) = u(a^k, y, x_1, x_2, x_3, x_4, x_5, a, b).

is true on this free semigroup \(S_2\).

G. S. Makanin [14] obtained a fundamental result in theory of equations in free semigroups (i.e., equations in words). Namely, he built an algorithm to recognize whether an arbitrary system of equations

\[w_i(x_1, \ldots, x_n, a_1, \ldots, a_m) = u_i(x_1, \ldots, x_n, a_1, \ldots, a_m)\]

is soluble in a free semigroup \(S_m\).

It follows immediately from this G. S. Makanin’s fundamental result that there exists an algorithm to recognize whether an arbitrary formula \(\Phi\) with quantifier prefix of the type \(\exists \exists \ldots \exists\)
or \( \forall \forall \ldots \forall \) (i. e. with the prefix containing only one type of quantifiers) is true on the semigroup \( S_m \).

Note that, given an arbitrary closed positive formula \( \Phi \) with the quantifier prefix of the type \( Q_1Q_2\ldots Q_n \forall \), it is possible to construct a closed positive formula \( \Phi^* \) with the quantifier prefix of the type \( Q_1Q_2\ldots Q_n \) such that formula \( \Phi \) is true on free semigroup \( S_m \) if and only if the formula \( \Phi^* \) is true on this semigroup.

Thus the question whether positive formulas with quantifier prefixes of the type \( (\exists x_1) \ldots (\exists x_n)(\forall y_1) \ldots (\forall y_m) \) are true on a free semigroup \( S_m \) is algorithmically decidable.

Another question is of interest as well: Does there exist an algorithm to recognize whether positive formulas with quantifier prefix of the type

\[
(\exists x_1)(\forall y)(\exists x_2)(\exists x_3)(\exists x_4)(\exists x_5) \text{ or } (\exists x_1)(\exists x_2)(\forall y)(\exists x_3)(\exists x_4)(\exists x_5)
\]

are true on a free semigroup \( S_2 \)?

Some authors consider two partial order relations, \( \leq \) and \( \subseteq \), on a semigroup \( S_m \) that are naturally defined as follows. For arbitrary elements \( X \) and \( Y \) of the semigroup \( S_m \)

\[
X \leq Y \iff \text{there exists an element } Z \text{ of the semigroup } S_m \text{ such that } Y = XZ;
\]

\[
X \subseteq Y \iff \text{there exists elements } U \text{ and } Z \text{ of the semigroup } S_m \text{ such that } Y = UXZ.
\]

This allows to consider formulas with bounded quantifiers of the type \( (Qz)_{z \leq t} \) and \( (Qz)_{z \subseteq t} \), where \( Q \) stands for \( \forall \) or \( \exists \), and \( t \) is a word in the variables and generators of the semigroup \( S_m \), which does not contain the variable \( z \).

### 3. Results and discussion

#### 3.1. Finitely generated free semigroups

Given an arbitrary finitely defined semigroup

\[
S = \langle a, b | A_1 = B_1, \ldots, A_n = B_n \rangle
\]

not containing empty defining words, let us build the formula \( \Phi^*_S(X, Y) \) of the type

\[
(\exists x)(\forall z)_{z \leq cxc}(\exists x_1)_{x_1 \subseteq cxc}y_{cxcYc}(\exists x_2)_{x_2 \subseteq cxcYc}(\exists x_3)_{x_3 \subseteq cxcYc}cXcxcYc = zax_1 \lor cXcxcYc = zb x_1 \lor cXcxcYc = zcy \lor \bigvee_{i=1}^{2^n} cXcxcYc = zcx_1A_i x_2 c x_1 B_i x_2 c x_3,
\]

where \( A_{n+j} = B_j \) and \( B_{n+j} = A_j \) for all \( 1 \leq j \leq n \).

**Lemma 1.** For arbitrary non-empty words \( A \) and \( B \) in the alphabet of generators of a semigroup \( S \), the following equivalence holds:

\[
A \text{ and } B \text{ define the same element of } S \iff \text{formula } \Phi^*_S(A, B) \text{ is true on the semigroup } S_3.
\]

**Proof.** Let us assume that the words \( A \) and \( B \) define the same element of the semigroup \( S \) (that is \( A = B \) in \( S \)). Then there exists such natural number \( m \) and such sequence of words \( W_0, W_1, \ldots, W_m \) of the free semigroup \( S_2 \) that

\[
W_0 = A, \quad W_m = B, \quad \text{and for all } 0 \leq i < m \text{ there exist such words } X_1 \text{ and } X_2 \text{ in } S_2\text{ that } W_i = X_1 A_i X_2 \text{ and } W_{i+1} = X_1 B_i X_2.
\]
For the sake of uniformity of the arguments, one can consider that \( m \geq 2 \). Assuming \( X_0 = W_1c\ldots cW_{m-1} \), it is easy to see that the following formula is true on the free semigroup \( S \):

\[
(\forall z)_{z \leq cAcX_0}(\exists x_1)_{x_1 \leq cAcX_0cBc}(\exists x_2)_{x_2 \leq cAcX_0cBc}(\exists x_3)_{x_3 \leq cAcX_0cBc}
\]

\[
\left( cAcX_0cBc = zax_1 \lor cAcX_0cBc = zbx_1 \lor cAcX_0cBc = zcBc \lor \bigvee_{i=1}^{2n} cAcX_0cBc = zcx_1A_i x_2 c x_1 B_i x_2 c x_3 \right). \tag{3}
\]

Therefore the formula \( \Phi_S(A, B) \) is true on the semigroup \( S_3 \) as well.

To prove the inverse, we assume that the formula \( \Phi_S(A, B) \) is true on the semigroup \( S_3 \). Let us chose such word \( X_0 \) in \( S_3 \) that the formula (3) is true on \( S_3 \).

There exist such natural number \( m \), such words \( W_m = A, W_m-1, \ldots, W_1, W_0 = B \) in the semigroup \( S_2 \), and such natural numbers \( \alpha_0 = 1, \alpha_1, \ldots, \alpha_m \) and \( \alpha_{m+1} = 1 \), that

\[
cAcX_0cBc = cW_m \alpha_m W_{m-1} \alpha_{m-1} W_{m-2} \alpha_{m-2} \ldots \alpha_2 W_1 \alpha_1 W_0 c.
\]

By induction on \( t \) we prove that \( W_t \) equals to \( W_0 \) in the semigroup \( S \). Indeed, let \( m \geq t > 0 \) and \( W_t \) equals to \( W_0 \) in the semigroup \( S \) for any \( t > i \geq 0 \):

We define the word \( Z_0 \) as follows

\[
Z_0 = \begin{cases} 
  cW_m \alpha_m W_{m-1} \alpha_{m-1} W_{m-2} \alpha_{m-2} \ldots \alpha_2 W_1 \alpha_1 W_0 c, & t < m, \\
  \text{empty word}, & t = m.
\end{cases}
\]

Then for some word \( U \) in the free semigroup \( S_3 \), the equality \( cAcX_0cBc = Z_0cU \) holds. Thus there exist such words \( X_1, X_2 \) and \( X_3 \) such that the following equality holds:

\[
cAcX_0cBc = Z_0cX_2cX_1B_iX_2cX_3 = Z_0cW_t \alpha_i \ldots \alpha_2 W_1 \alpha_1 W_0 c.
\]

The following three cases are possible.

I. Both the words \( X_1 \) and \( X_2 \) do not contain the letter \( c \) (i.e. \( X_1 \) and \( X_2 \) are elements of the semigroup \( S_2 \)). Then

\[
W_t = X_1 A_i X_2 \text{ and } W_{t-1} = X_1 B_i X_2.
\]

Thus \( W_t \) equals to \( W_{t-1} \) in the semigroup \( S \), but \( W_{t-1} \) equals to \( W_0 \) by the inductive hypothesis. Therefore \( W_t \) equals to \( W_0 \) in the semigroup \( S \).

II. The word \( X_1 \) contains the letter \( c \). Then there exists such word \( X_{1r} \) in the semigroup \( S_3 \), that

\[
cX_1 = cW_i cX_{1r}.
\]

Therefore there exists \( j < t \) such that \( W_t = W_j \). By the inductive hypothesis, \( W_j \) equals to \( W_0 \) in the semigroup \( S \). Therefore \( W_t \) equals to \( W_0 \) in the semigroup \( S \).

III. The letter \( c \) does not appear in the word \( X_1 \), but do appear in the word \( X_2 \). In this case there exists such word \( X_{2r} \) in the semigroup \( S_2 \) and such word \( X_{2r} \) in the semigroup \( S_3 \), that

\[
X_2 = X_2 cX_{2r}, \quad cW_i c = cX_1 A_i X_2 c \quad W_i = X_1 A_i X_2 c.
\]

There exists \( j < t \) such that \( cX_1 B_i X_{2r} = cW_j c \). Therefore \( X_1 B_i X_{2r} = W_j \). By the inductive hypothesis, \( W_j \) equals to \( W_0 \) in the semigroup \( S \). Therefore \( W_t \) equals to \( W_0 \) in the semigroup \( S \).
We use the notation and the results of Ref. [12] to remove the disjunction symbol \( \lor \) from the formula (2). For an arbitrary word \( w \) we use the notation [12]: \( \langle w \rangle = wawb \). The following equivalence is proven in Ref. [12] for an arbitrary semigroup \( S \): 

\[
\bigvee_{i=1}^{n} W = W_i \iff (\exists Z)(\exists Z') U = ZVZ',
\]

where \( v = WW_1 \ldots W_n, \quad V = \langle v \rangle^2 \langle v \rangle^2, \quad U = \langle v \rangle^2 W_1 \langle v \rangle^2 W_2 \langle v \rangle^2 \ldots \langle v \rangle^2 W_n \langle v \rangle^2 \).

It is easy to see that \( Z, Z' \subseteq U \). Given the formula \( \Phi_S(X,Y) \), this fact allows to build the formula \( \Phi_S'(X,Y) \) of the type

\[
(\exists x)(\forall z)_{z \leq t}(\exists x_1)_{x_1 \leq t_1}(\exists x_2)_{x_2 \leq t_1}(\exists x_3)_{x_3 \leq t_1}(\exists x_4)_{x_4 \leq t_2}(\exists x_5)_{x_5 \leq t_2} w = v,
\]

where \( t = cXcx, t_1 = cXcxcYc, t_2 = U \)

such that the following equivalence holds for arbitrary non-empty words \( A \) and \( B \) in the alphabet of the generators of the semigroup \( S \):

words \( A \) and \( B \) define the same element of the semigroup \( S \):

\[
\iff \quad \text{formula } \Phi_S(A,B) \text{ is true on the semigroup } S_3.
\]

If we assume that (i) \( S \) is a semigroup with non-empty relation words, (ii) the problem of identity to the non-empty word \( B \) is algorithmically undecidable in \( S \) and (iii) \( \Phi_S(X,Y) \) stands for the formula \( \Phi_S(X) \), then we obtain the following theorem.

**Theorem 3.** Under the above assumptions (i)–(iii), it is possible to construct a family of formulas \( \Phi_S(X) \) with one parameter \( X \) of the type

\[
(\exists x)(\forall z)_{z \leq t}(\exists x_1)_{x_1 \leq t_1}(\exists x_2)_{x_2 \leq t_1}(\exists x_3)_{x_3 \leq t_1}(\exists x_4)_{x_4 \leq t_2}(\exists x_5)_{x_5 \leq t_2} w(X, x, z, x_1, x_2, x_3, x_4, x_5, a, b, c) = u(X, x, z, x_1, x_2, x_3, x_4, x_5, a, b, c),
\]

such that there is no algorithm to recognize, given an arbitrary word \( A \) of the free semigroup \( S_2 \), whether the positive formula \( \Phi_S(A) \) is true on the free semigroup \( S_3 \).

We remark that the formulas under consideration contain only one unbounded quantifier, \( \exists \). As for the formulas that contain bounded quantifiers only (the quantifier-free part being arbitrary), the problem, whether these formulas are true on an arbitrary free semigroup \( S_m \), is algorithmically decidable.

### 3.2. Countably generated free semigroups

The finitely generated free semigroups are somewhat different from the finitely generated free semigroups considered in the previous Section. A. D. Taimanov and Ju. I. Khmelevskii [15] showed that the universal theory of a free semigroup \( S_\omega \) of countable rank is algorithmically decidable, i. e. that there exists an algorithm to recognize, given an arbitrary formula of the type

\[
(\forall x_1)(\forall x_2)\ldots(\forall x_n) \Psi,
\]

where \( \Psi = \bigvee_{i=1}^{k} \left( (\&_{j \in A_i} w_{ij} = u_{ij}) \& (\&_{t \in B_i} v_{it} \neq z_{it}) \right), \)

\( w_{ij}, u_{ij}, v_{it}, z_{it} \) are words in the alphabet \( \{ x_1, \ldots, x_n, a_1, \ldots, a_m, \ldots \} \),

\( A_i \) and \( B_i \) are some sets (possible empty),
whether it is true on the free semigroup $S_\omega$. Clearly, this also holds for the formulas of the type

$$(\exists x_1)(\exists x_2)\ldots(\exists x_n)\, \Psi,$$

where $\Psi = \bigvee_{i=1}^k \left( (\wedge_{j \in A_i} w_{ij} = u_{ij}) \& (\wedge_{t \in B_i} v_t \neq z_t) \right)$,

$w_{ij}, u_{ij}, v_t, z_t$ are words in the alphabet $\{ x_1, \ldots, x_n, a_1, \ldots, a_m, \ldots \}$.

$A_i$ and $B_i$ are some sets (possible empty),

The counterpart of this result for finitely generated groups clearly follows from the fundamental theorem by G. S. Makanin [14]. Indeed, as was mentioned above, the relation $x \neq y$ can be expressed by a positive $\exists$-formula. However, such an approach does not work for a countably generated semigroup, so A. D. Taimanov and Ju. I. Khmelevskii used another method. Nevertheless, the G. S. Makanin’s theorem can be used if we note that for the formula $\Phi$ of the type

$$(\forall x_1)(\forall x_2)\ldots(\forall x_n)\, \Psi,$$

where $\Psi = \bigvee_{i=1}^k \left( (\wedge_{j \in A_i} w_{ij} = u_{ij}) \& (\wedge_{t \in B_i} v_t \neq z_t) \right)$,

where $w_{ij}, u_{ij}, v_t, z_t$ – are words in the alphabet $\{ x_1, \ldots, x_n, a_1, \ldots, a_m \}$,

$A_i$ and $B_i$ – are some sets (possible empty),

the following equivalence holds:

$$S_\omega \models \Phi \iff S_{m+2} \models \Phi.$$ 

If $S_\omega \models \Phi$, then clearly $S_{m+2} \models \Phi$ as well.

To prove the inverse, let us assume that $S_{m+2} \models \Phi$. Then the formula $\Phi$ is true on the countably generated sub-semigroup of $S_{m+2}$:

$$\langle a_1, \ldots, a_m, a_{m+1}a_{m+2}a_{m+1}, a_{m+1}a_{m+2}^2a_{m+1}, \ldots, a_{m+1}a_{m+2}^n a_{m+1}, \ldots \rangle.$$ 

The map $\varphi$ defined by

$$\begin{cases}
\varphi(a_n) = a_n & \text{where } n = 1, 2, \ldots, m; \\
\varphi(a_{m+1}a_{m+2}^{n-m}a_{m+1}) & \text{where } n = m + 1, m + 2, \ldots
\end{cases}$$

is an isomorphism of the free group $S_\omega$ onto this countably generated sub-semigroup, therefore $S_\omega \models \Phi$.

It was noted in Ref. [16] that the positive theory of free semigroup $S_\omega$ is recursively enumerable and the question whether it is recursive (i.e. algorithmically decidable) was addressed.

To prove that the positive theory of a free semigroup $S_\omega$ of countable rank is recursively enumerable, one can use the technique applied by Yu. I. Merzlyakov [17] to free groups:

given an arbitrary formula $\Phi$, which contains only the free generators, $a_1, a_2, \ldots, a_m$, it is possible to find an equation with constraints of the type

$$w(x_1, \ldots, x_q, a_1, \ldots, a_m, a_{m+1}, \ldots, a_p) = u(x_1, \ldots, x_q, a_1, \ldots, a_m, a_{m+1}, \ldots, a_p) \&$$

$$\& \quad \forall_{i=1}^q x_i \in \langle a_1, \ldots, a_{n_i} \rangle,$$

where $\langle a_1, \ldots, a_{n_i} \rangle$ is the sub-semigroup generated by the elements $a_1, \ldots, a_{n_i}$, $m \leq n_1 \leq n_2 \leq \ldots \leq n_q = p$, such that formula $\Phi$ is true on the free semigroup $S_\omega$ iff the above equation with constraints is solvable in the free semigroup $S_p$. 

8
Having made some modifications of the G. S. Makanin’s algorithm, Yu. M. Vazhenin and B. V. Rozenblat [18] built an algorithm to recognize solvability of an arbitrary system of equations with the constraints of the above type in a free semigroup. Thus Yu. M. Vazhenin and B. V. Rozenblat proved decidability of positive theory (with constants) of free semigroup of countable rank. By a similar way, using the above-mentioned Yu. I. Merzlyakov’s results [17], G. S. Makanin proved [19] decidability of positive theory (with constants) of free group of countable rank.

The results by G. S. Makanin’, Yu. M. Vazhenin and B. V. Rozenblat were strengthen by K. U. Schulz [20] who modified the G. S. Makanin’s algorithm to recognize, given an arbitrary system of equations with the constraints of the type

\[
\frac{p}{\&} \ w_i(x_1, \ldots, x_n, a_1, \ldots, a_m) = u_i(x_1, \ldots, x_n, a_1, \ldots, a_m) \& \frac{n}{\&} x_i \in L_i,
\]

where \(L_1, \ldots, L_n\) are regular languages, whether this system is solvable in the free group \(S_m\). We remark here that any finitely generated semigroup is a regular set.

Given an arbitrary finitely defined semigroup

\[S = \langle a, b | A_1 = B_1, \ldots, A_n = B_n \rangle\]

not containing empty relation words, let us build the formula \(\Phi_S(X, Y)\) of the type

\[
(\exists x)(\forall y)(\forall z)(\exists x_1)(\exists x_2)(\exists x_3)((cXcx \neq ycaz \& cXcx \neq ycbz) \lor \\
\lor \bigvee_{i=1}^{2n} cXcxYc = ycx_1A_i x_2cX_1B_i x_2cX_3),
\]

where for any \(1 \leq j \leq n: A_{n+j} = B_j\) and \(B_{n+j} = A_j\) (symmerization of the set of relators).

To clarify the further considerations we note that the above formula \(\Phi_S(X, Y)\) is equivalent to the formula of the type

\[
(\exists x)(\forall y)(\forall z)(\exists x_1)(\exists x_2)(\exists x_3)((cXcx = ycaz \lor cXcx = ycbz) \rightarrow \\
\rightarrow \bigvee_{i=1}^{2n} cXcxYc = ycx_1A_i x_2cX_1B_i x_2cX_3).\]

**Lemma 2.** For arbitrary non-empty words \(A\) and \(B\) in the alphabet of generators of the semigroup \(S\), the following equivalence holds:

the words \(A\) and \(B\) define the same element of the semigroup \(S\)

\[\iff \text{formula } \Phi_S(A, B) \text{ is true on the semigroup } S_w.\]

**Proof.** Let the words \(A\) and \(B\) define the same element of the semigroup \(S\) (i. e. equal in the semigroup \(S\)). Then there exists such natural number \(m\) and such sequence of words \(W_0, W_1, \ldots, W_m\) in the free semigroup \(S_2\), that

\[W_0 = A, \quad W_m = B \quad \text{and for all } 0 \leq i < m \text{ there exist such words } X_1 \text{ and } X_2, \text{ in } S_2, \text{ that } W_i = X_1 A_i X_2 \text{ and } W_{i+1} = X_1 B_i X_2.\]

For the sake of uniformity of the considerations, one can assume that \(m \geq 2\).

Let \(X_0 = W_1 c \ldots c W_{m-1}\). It is easy to see that the formula

\[
(\forall y)(\forall z)(\exists x_1)(\exists x_2)(\exists x_3)((cAcX_0 \neq ycaz \& cAcX_0 \neq ycbz) \lor \\
\lor \bigvee_{i=1}^{2n} cAcX_0Bc = ycx_1A_i x_2cX_1B_i x_2cX_3),
\]

is true on the semigroup \(S_w\).
is true on the free semigroup $S_\omega$. Therefore, the formula $\Phi_S(A, B)$ is true on the semigroup $S_\omega$.

To prove the inverse, we assume the formula $\Phi_S(A, B)$ to be true on the semigroup $S_\omega$. In the semigroup $S_\omega$ we choose such word $X_0$ (of minimal possible length) that, the formula

$$(\forall y)(\forall z)(\exists x_1)(\exists x_2)(\exists x_3)((cAcX_0 \neq ycaz \& cAcX_0 \neq ycbz) \vee$$

$$\vee \bigvee_{i=1}^{2n} cAcX_0cBc = ycxc_1A_ixc_2xc_1B_ixc_2xc_3)_i,$$

is true on $S_\omega$. It is easy to see that the word $X_0$ cannot be empty.

For the semigroup $S$, let us build the chain of elementary transformations

$$A = W_0 \rightarrow W_1 \rightarrow \ldots \rightarrow W_m = B,$$

which transform the word $A$ to the word $B$.

Let $W_0 = A$.

Taking the empty word as $y$, we obtain that for some words $X_1$, $X_2$, $X_3$ and some $j$ ($1 \leq j \leq 2n$) the following equality holds:

$$cAcX_0cBc = cX_1A_jx_2cX_1B_jx_2cX_3.$$

If the word $X_1$ contains letters other than $a$ and $b$, then the word $A$ is the beginning of the word $X_1$, and for some word $X'_0$ the equality

$$cAcX'_0cBc = cX_1B_jx_2cX_3$$

holds and the formula

$$(\forall y)(\forall z)(\exists x_1)(\exists x_2)(\exists x_3)((cAcX'_0 \neq ycaz \& cAcX_0 \neq ycbz) \vee$$

$$\vee \bigvee_{i=1}^{2n} cAcX_0cBc = ycxc_1A_ixc_2xc_1B_ixc_2xc_3),$$

is true, that contradicts the above assumption that $X_0$ has the minimal possible length. Thus $X_0$ does not contain the letter other than $a$ and $b$. But then, for some words $X''_2$ and $X''_2$, the following equalities hold:

$$X_2 = X_2''X_2'' \& A = X_1A_jX_2'.$$

Let $W_1 = X_1B_jX_2'$.

To build the word $W_2$ we take the word $cX_1A_jX_2$ as $y$, and the similar considerations give the transformation

$$X_1B_jX_2' = W_1 \rightarrow W_2.$$

As a result, for the semigroup $S$ we obtain the chain of elementary transformations:

$$A = W_0 \rightarrow W_1 \rightarrow \ldots \rightarrow W_m = B,$$

which transform the word $A$ to the word $B$.

As a result, for the semigroup $S$ we obtain the chain of elementary transformations:

$$A = W_0 \rightarrow W_1 \rightarrow \ldots \rightarrow W_m = B,$$

which transform the word $A$ to the word $B$.

Let the semigroup $S$ be a semigroup with non-empty relation words, and let the problem of equality to an non-empty word $B$ is algorithmically undecidable in $S$. Then the following theorem holds:

**Theorem 4.** The elementary $\exists \forall^2 \exists^3$-theory of free semigroup $S_\omega$ of countable rank is algorithmically undecidable.
The formula $\Phi_S(A, B)$, which appears in the above proofs, can be expressed in the form:

$$(\exists x)(\forall y_1)(\forall y_2)(\forall y_3)(\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4)(\exists x_5)(u \neq v \lor w_1 = w_2),$$

i.e. in the form:

$$(\exists x)(\forall y_1)(\forall y_2)(\forall y_3)(\exists x_1)(\exists x_2)(\exists x_3)(\exists x_4)(\exists x_5)(u \rightarrow v \rightarrow w_1 = w_2).$$

Given an arbitrary finitely defined semigroup

$$S = \langle a, b | A_1 = B_1, \ldots, A_n = B_n \rangle$$

which does not contain empty relation words, let us build the formula $\Phi_S^\omega(X, Y)$ of the type

$$(\exists x)(\forall z)zc \leq x \exists (\exists x_1)x_1 \subseteq cXcXcYc(\exists x_2)x_2 \subseteq cXcXcYc(\exists x_3)x_3 \subseteq cXcXcYc$$

$$(\forall t cXcXcYc = zcx_1A_1x_2cXcB_1x_2cXc),$$

where for all $1 \leq j \leq n$: $A_{n+j} = B_j$ and $B_{n+j} = A_j$.

The following lemma can be proved using the technique similar to that used to prove the lemmas 1 and 2.

**Lemma 3.** For arbitrary non-empty words $A$ and $B$ in the alphabet of generators of a semigroup $S$, the following equivalence holds:

words $A$ and $B$ define the same element of the semigroup $S$

$$\iff$$

formula $\Phi_S^\omega(A, B)$ is true on the semigroup $S_\omega$.

Using the above technique, we remove the disjunction sign $\lor$ from the formula $\Phi_S^\omega(A, B)$. This results in the formula $\Phi_S^\omega(X, Y)$ of the type

$$(\exists x)(\forall z)zc \leq t(\exists x_1)x_1 \subseteq t_1(\exists x_2)x_2 \subseteq t_1(\exists x_3)x_3 \subseteq t_1(\exists x_4)x_4 \subseteq t_2(\exists x_5)x_5 \subseteq t_2 w = v,$$

where $t = cXcX$, $t_1 = cXcXcYc$, $t_2 = U$

such that for any non-empty words, $A$ and $B$, in the alphabet of generators of the semigroup $S$, the following equivalence holds:

words $A$ and $B$ define the same element of the semigroup $S$

$$\iff$$

formula $\Phi_S^\omega(A, B)$ is true on the semigroup $S_\omega$.

Let the semigroup $S$ be a semigroup with non-empty relation words, and let the problem of equality to an non-empty word $B$ is algorithmically undecidable in $S$. Then using the formula $\Phi_S^\omega(X, B)$ instead of the formula $\Phi_S^\omega(X)$, we obtain the following

**Theorem 5.** One can build a one-parameter family of formulas $\Phi_S^\omega(X)$ with the parameter $X$ of the type

$$(\exists x)(\forall z)zc \leq t(\exists x_1)x_1 \subseteq t_1(\exists x_2)x_2 \subseteq t_1(\exists x_3)x_3 \subseteq t_1(\exists x_4)x_4 \subseteq t_2(\exists x_5)x_5 \subseteq t_2$$

$$w(X, x, z, x_1, x_2, x_3, x_4, x_5, a, b, c) = u(X, x, z, x_1, x_2, x_3, x_4, x_5, a, b, c),$$

such that there does not exist an algorithm to recognize, given an arbitrary word $A$ (an element of the free semigroup $S_2$), whether the positive formula $\Phi_S^\omega(A)$ is true on the free semigroup $S_\omega$.  

11
4. Conclusion
Recognizing the algorithmical (un)decidability is important for wide variety of problems in word combinatorics – from number and coding theories to theoretical informatics and computer science. We demonstrate some “simple” fragments of positive theories of finitely generated semigroups and elementary theory of a free countably generated semigroup and show that they are algorithmic undecidable. The “simplicity” relates to the quantifier-prefix and the quantifier-free part of the formulas, and means that all the above considered formulas have only one unbounded quantifier (existence quantifier $\exists$). If a formula (with arbitrary quantifier-free part) contains bound quantifiers only, the problem of its truth on an arbitrary free semigroup $S_\omega$ is algorithmically decidable.

References
[1] Diekert V 2015 More than 1700 years of word equations Algebraic Informatics ed Maletti A (Cham: Springer International Publishing) pp 22–28 ISBN 978-3-319-23021-4
[2] Lothaire M 2005 Applied Combinatorics on Words Encyclopedia of Mathematics and its Applications (Cambridge: Cambridge University Press) ISBN 978-1-107-34100-5
[3] Freydenberger D D 2019 Theory Comput. Syst. 63 1679–1754 ISSN 1433-0490
[4] Zheng Y, Ganesh V, Subramanian S, Tripp O, Berzish M, Dolby J and Zhang X 2017 Form. Methods. Syst. Des. 50 249–288 ISSN 1572-8102
[5] Day J D, Ganesh V, He P, Manea F and Nowotka D 2018 The satisfiability of word equations: Decidable and undecidable theories Reachability Problems ed Potapov I and Reynier P A (Cham: Springer International Publishing) pp 15–29 ISBN 978-3-030-00250-3
[6] Diekert V, Jeż A and Plandowski W 2016 Inform. Comput. 251 263–286 ISSN 0890-5401
[7] Halfon S, Schoenebelen P and Zetzsche G 2017 Decidability, complexity, and expressiveness of first-order logic over the subword ordering 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS) (Reykjavik, Iceland) pp 1–12
[8] Quine W V 1946 J. Symbol. Logic 11 105–114
[9] Durnev V G 1973 Dokl. Akad. Nauk SSSR 211(4) 772–774 in Russian
[10] Kosovskiy N K 1981 Elements of Mathematical Logic and its Application to the theory of Subrecursive Algorithms (Leningrad: LSU Publ.) in Russian
[11] Durnev V G 1995 Siberian Math. J. 36 917–929 ISSN 1573-9260
[12] Karhumäki J, Plandowski W and Mignosi F 1997 The expressibility of languages and relations by word equations Automata, Languages and Programming ed Degano P, Gorrieri R and Marchetti-Spaccamela A (Berlin, Heidelberg: Springer) pp 98–109 ISBN 978-3-540-69194-5
[13] Marchenkov S S 1982 Siberian Math. J. 23 196–198 in Russian
[14] Makanin G S 1977 Math. USSR-Sb. 32 129–198
[15] Taimanov A D and Hmelevskii J I 1980 Sibirsk. Mat. Z. 21 228–230 in Russian
[16] Durnev V G 1972 The positive theory of a free semigroup Questions in the Theory of Groups and Semigroups ed Grindlinger M (Tula: Tul’sk. Gos. Ped. Inst.) pp 122–172 in Russian
[17] Merzlyakov Y I 1966 Algebra i logika 5 25–42 in Russian
[18] Vazhenin Y M and Rozenblat B V 1983 Math. USSR-Sb. 44(1) 109–116
[19] Makanin G S 1985 Math. USSR Izv. 25 75–88
[20] Schulz K U 1992 Makanin’s algorithm for word equations-two improvements and a generalization Word Equations and Related Topics ed Schulz K U (Berlin, Heidelberg: Springer) pp 85–150 ISBN 978-3-540-46737-3