# Word Problem for Thue Systems with a Few Relations

Yuri Matiyasevich

Steklov Institute of Mathematics
of Russian Academy of Sciences
Saint-Petersburg Branch (POMI RAN)
27 Fontanka
Saint-Petersburg
191011, Russia

**Abstract.** The history of investigations on the word problem for Thue systems is presented with the emphasis on undecidable systems with a few relations. The best known result, a Thue system with only three relations and undecidable word problem, is presented with details. Bibl. 43 items.

## 1 Historical Introduction

The *general* theory of rewriting systems is a relatively new area. However, *particular* kinds of rewriting systems were known and studied in mathematics for a long time, in particular, in the theory of semigroups.

Let $S = \langle M, \circ \rangle$ be a *semigroup*, i.e., a set $M$ and a binary operation on it satisfying the associative law

$$(m_1 \circ m_2) \circ m_3 = m_1 \circ (m_2 \circ m_3)$$

for any $m_1$, $m_2$, $m_3$ from $M$. This law allows one to omit parentheses and write

$$m_1 \circ m_2 \circ m_3$$

or simply

$$m_1 m_2 m_3$$

without ambiguity.

For a semigroup $S = \langle M, \circ \rangle$ there may exist a finite subset

$$A = \{a_1, \ldots, a_n\} \subseteq M$$

such that any element of $M$ can be constructed from $a_1, \ldots, a_n$ with the aid of the operation $\circ$, i.e., every element $m$ from $M$ has a *representation*

$$m = a_{i_1} \circ \ldots \circ a_{i_k} \tag{1}$$

where $1 \leq i_j \leq n$. In such a case we say that $S$ is *finitely generated* and $A$ is a *set of generators*. If each $m$ from $M$ has only one representation (1), then the semigroup is called *free* and, treating $A$ as an alphabet, we can identify $M$ with $A^*$, the set of all words in $A$, and identify $\circ$ with concatenation.

More interesting is the case when some elements have more than one representation and hence we have non-trivial *relations*

$$a_{i_1} \dots a_{i_l} = a_{j_1} \dots a_{j_l} \tag{2}$$

among the generators.

As soon as we have a single relation, we have countably many of them because (2) implies that

$$\begin{gathered} ma_{i_1} \dots a_{i_l} = ma_{j_1} \dots a_{j_l}, \\ a_{i_1} \dots a_{i_l} m = a_{j_1} \dots a_{j_l}\ m \end{gathered}$$

for any $m$ from $M$. Also if we have two relations

$$P = Q$$

and

$$Q = R$$

then we have relations

$$Q = P$$

and

$$P = R$$

as well.

It may happen that there is a finite set

$$\begin{gathered} P_1 = Q_1, \\ \vdots \\ P_t = Q_t \end{gathered} \tag{3}$$

of relations such that any other relation $P = Q$ is implied by relations (3) in the above described way. In such a case the semigroup is called *finitely presented.* A *presentation*, consisting of the alphabet and the set of relations, is also called a *Thue system* and can be viewed as a special kind of rewriting system.

Relation are often written as

$$\begin{gathered} P_1 \longleftrightarrow Q_1, \\ \vdots \\ P_t \longleftrightarrow Q_t. \end{gathered} \tag{4}$$

Two words, $G$ and $H$, from $A^*$ are said to be *immediately equivalent* in a given Thue system $T$, written $G \underset{T}{\longleftrightarrow} H$, if

- either they are graphically equal, written $G \equiv H$ (i.e., equal as words in $A^*$)
- or there are number $i$ and words $X$ and $Y$ such that $1 \leq i \leq t$ and
  - either $G \equiv XP_iY$ and $H \equiv XQ_iY$
  - or $G \equiv XQ_iY$ and $H \equiv XP_iY$.

(To make the further exposition a bit smoother the author has deviated from the traditional definition of immediate equivalence; usually this relation is not required be reflexive.)

The transitive closure of $\underset{T}{\longleftrightarrow}$ is denoted by $\overset{*}{\underset{T}{\longleftrightarrow}}$: two words $G$ and $H$ are *equivalent* in $T$, written $G \overset{*}{\underset{T}{\longleftrightarrow}} H$, if there are words $W_0, \ldots, W_k$ which provide a *derivation*

$$G \equiv W_0 \underset{T}{\longleftrightarrow} W_1 \underset{T}{\longleftrightarrow} \ldots \underset{T}{\longleftrightarrow} W_k \equiv H$$

of word $H$ from word $G$.

The choice of generators is not unique so a finitely presented semigroup has in general many presentation. On the other hand, given Thue system determines corresponding semigroup up to isomorphism. Namely, one can identify $M$, the set of elements, with the set of classes of equivalent words; if $\mathcal{G}$ and $\mathcal{H}$ are two such classes and $G_1, G_2 \in \mathcal{G}$, $H_1, H_2 \in \mathcal{H}$, then words $G_1H_1$ and $G_2H_2$ are also equivalent and $\mathcal{G} \circ \mathcal{H}$ can be defined as the unique class containing these words.

The name "Thue system" was given after Axel Thue who in 1914 posed in [37] the following problem:

> *Given a finite presentation of a semigroup and two words in the alphabet of generators, determine whether the words are equivalent.*

This problem is known as *Thue's problem* or *word problem for finitely presented semigroups.* One can equally speak about finitely pesented monoid because the empty word, denoted by $\Lambda$, evidently plays the role of unit under concatenation. In the Russian literature Thue systems are also known under the name *associative systems* or *associative calculi* and the word problem is also called *the equivalence problem.*

Similar problem was posed a few years earlier by Dehn [12, 13] about groups. A *finitely presented group* can be defined as a finitely presented semigroup in which for every generator $a$ there is another generator, denoted by $a^{-1}$, and two relations

$$\begin{aligned} aa^{-1} &= \Lambda, \\ a^{-1}a &= \Lambda. \end{aligned} \tag{5}$$

When counting the number of generators, $a$ and $a^{-1}$ are counted for one, and *trivial* relations (5) are not included into the count of relations in a presentation of a group. Corresponding problem for groups is known as *Dehn's problem* or *word problem for finitely presented groups.*

Dehn's problem for groups with single defining relation was solved in 1932 by Magnus [20]. The reader should realize that even at that time (to say nothing about the beginning of the century when the problems were posed) there were neither Turing machines no partial recursive functions, Church's Thesis was not stated yet and the mankind lived in happy ignorance of algorithmically unsolvable problems. Dehn and Thue posed their problems in positive sense: *find* a method which would allow to determing whether the words are equivalent or not.

The situation changed in the middle of 30's with the development of rigorous *general* notion of algorithm and Church's Thesis. These achievements, together with the first examples of algorithmically undecidable problems, formed a basis for tackling Dehn's and Thue's problems in the negative direction.

With respect to semigroups the success was achieved in 1947 by Markoff [23] and by Post [33]. The algorithmical undecidability of Thue's problem was established (see [24]) in the following strong sense:

> *there is a Thue system $T$ and a word $G$ such that there is no algorithm to decide, given a word $H$, whether it is equivalent to $G$ or not.*

Thue's problem was the first decision problem which arose in mathematics proper (i.e., not in logic or calculability theory) and which was shown algorithmically undecidable.

The above cited result by Magnus indicated that a semigroup with undecidable word problem should have sufficiently many defining relations and one could hope to find partial positive solution of Thue's problem for semigroups defined by a small number of relations. It was natural to seek for the boundary between decidable and undecidable in terms of the number of defining relations, and several researchers devoted their investigations to constructing examples of semigroups with undecidable word problem and a small number of defining relation.

Markoff constructed such a semigroup with 13 generators and 33 relations. It was published at first in short note [24] and later detailed proof was supplied in [25]; see also Lallement's exposition [18] in this volume.

Markoff's result was improved several years later by Scott [35] and by Tseitin to 7 relations. This drastic cutting of the number of relations became possible thanks to Dehn's problem being at that time proved undecidable (both Scott and Tseitin based their constructions on Novikov's result [29, 30]; today examples of finitely presented groups with undecidable word problem can be found in [3, 16, 7, 6, 42, 40, 11, 10]). Tseitin's example was published at first in short note [38] and detailed proof was given in [39]; see also Lallement's exposition [18] in this volume.

Tseitin's semigroup is undecidable in the above mentioned strong sence: *there is a particular word $G$ such that there is no algorithm to decide, given a word $H$, whether it is equivalent to $G$ or not.* While this semigroup has only 7 relations containing 33 occurences of 5 generators, the length of such a word $G$ known today is measured in thousands of letters. That is why Tseitin has also constructed another Thue system with 9 relations

$$\begin{aligned} ac &\longleftrightarrow ca, \\ ad &\longleftrightarrow da, \\ bc &\longleftrightarrow cb, \\ bd &\longleftrightarrow db, \\ eca &\longleftrightarrow ce, \\ edb &\longleftrightarrow de, \\ cdca &\longleftrightarrow cdcae, \\ caaa &\longleftrightarrow aaa, \\ daaa &\longleftrightarrow aaa \end{aligned}$$

and proved that *is no algorithm to decide, given a word $H$, whether it is equivalent to word $aaa$ or not.*

Further progress to undecidable Thue systems with 5 relations was achieved by Makanin [21] and by Matiyasevich [26, 27] as modification of Tseitin's example with 7 relations.

In 1967 the author [26] constructed a Thue system

$$\begin{aligned} \alpha\alpha\sigma\alpha\sigma &\longleftrightarrow \sigma\alpha\alpha, \\ \alpha\alpha\sigma\sigma &\longleftrightarrow \sigma\alpha\alpha, \\ L &\longleftrightarrow M \end{aligned} \tag{6}$$

in two letter alphabet $\{\alpha, \sigma\}$ for which the word problem is undecidable. The number of relations, 3, still remains the record one. The cost paid for this reduction of the number of relations was the length of relations, namely word $L$ had 304 letters and word $M$ had 608 letters, so Tseitin's example with 7 relations remains the shortest one, i.e., the one with the least total number of letters in the presentation.

Surprisingly, we still do not have a counterpart of Magnus's result for semigroups with one defining relation. Only partial progress for single relation of special forms was achieved (see surveys by Adian and Makanin [2] and by Lallement [18] in this volume; also see [43]).

Parallel to reduction of the number of defining relations in undecidable semigroups there was reduction of the number of relations in undecidable groups. Here there were two sources of progress. On one hand, there were new ideas specific for the case of groups. On the other hand, constructions of undecidable groups were based on undecidable semigroups, so a reduction for semigroups immediately caused corresponding reduction for groups. Today the smallest number, 12, is achieved in the group constructed by Borisov [6] on the base of authour's semigroup (6) with 3 relations.

Besides groups, semigroup (6) was used for constructing "simple" examples of other undecidable problem. A Thue system (4) with $t$ relations can be rewritten as *semiThue system*

$$\begin{aligned} P_1 &\longrightarrow Q_1, \\ Q_1 &\longrightarrow P_1, \\ &\vdots \\ P_t &\longrightarrow Q_t, \\ Q_t &\longrightarrow P_t \end{aligned} \tag{7}$$

with $2t$ *rules* (relations $\xrightarrow[T]{}$ and $\xrightarrow[T]{*}$ are defined similar to the above definitions of $\xleftrightarrow[T]{}$ and $\xleftrightarrow[T]{*}$ with the last possibility, $G \equiv XQ_iY$ and $H \equiv XP_iY$, being omitted). Claus [8, 9] (see also [14]) proved that a semiThue system with undecidable word problem and $k$ rules enable one to prove undecidability of *Post correspondence problem* introduced in [33] for $k+4$ pairs of words. Using (6), Claus established undecidability of Post correspondence problem for 10 pairs of words. Pansiot noticed in [31] that the first two relations in (6) had equal right-hand side parts and hence could be replaced by 3 semiThue rules

$$\sigma\alpha\alpha \longrightarrow \alpha\alpha\sigma\alpha\sigma,$$

$$\alpha\alpha\sigma\alpha\sigma \longrightarrow \alpha\alpha\sigma\sigma,$$
$$\alpha\alpha\sigma\sigma \longrightarrow \sigma\alpha\alpha,$$

which together with two rules $L \longrightarrow M$ and $M \longrightarrow L$ form an undecidable semiThue system with 5 rules and, according to Claus's result, imply the undecidability of Post correspondence problem for 9 pairs of words.

The undecidability of the word problem for a semiThue system with 5 rules and the undecidability of Post correspondence problem for 9 pairs of words remained the best results until recently. In December 1993, Sénizergues and the author were able to construct a semiThue system with 4 rules only and with undecidable word problem, which respectively implied the undecidability of Post correspondence problem for 8 pairs of words. Besides the old main idea used for construction of Thue system (6), this new result required a number of new ideas specific to semiThue systems.

The undecidable Thue system (6) has got wide citation (see [1–2, 4–6, 8–11, 14, 17–19, 22, 28, 31, 34, 36, 40–43]) in spite of the fact that the author had never made his proof of undecidability available to broad readership. The only place where the detailed proof had been given by the author was his thesis existing in a few copies. One excuse for this was prompt appearence of Collins's paper [11] containing a detailed proof. However, this proof was rather different from author's original one. An idea of this original proof can be obtained from a joint publication of Boone, Collins and Matiyasevich [4] where this technique was used for obtaining similar but a bit different result; however, this publication is not easily available.

It gives a pleasure to the author to take the opportunity providied by this *Ecole de Printemps* and publish his original proof. The particular Thue system (6) presented in [26] was based on Tseitin's example and hence eventually depended on the deep and difficult result about the unsolvability of Dehn's problem for groups. Below the technique is introduced in a general setting which enables one to construct a Thue system with 3 relation and undecicable word problem starting from any Thue system for which this problem is undecidable.

The author expresses his gratitude to Géraud Sénizergues who read carefully the manuscript and revealed numerous bugs.

## 2 Construction

### 2.1 Plan

Let $T_0$ be an arbitrary Thue system

$$\begin{array}{c} R_1 \longleftrightarrow S_1, \\ \vdots \\ R_{t_0} \longleftrightarrow S_{t_0} \end{array} \tag{8}$$

in some alphabet $A_0 = \{a_1, \ldots, a_n\}$. We shall construct another Thue system $T_3$ defined by 3 relation in a two letter alphabet $A_3$, and a mapping

$$\sigma : A_0^* \to A_3^*$$

such that two words $G_0$ and $H_0$ from $A_0^*$ are equivalent in $T_0$ if and only if their images, $\sigma(G_0)$ and $\sigma(H_0)$, are equivalent in $T_3$, i.e.,

$$G_0 \xleftrightarrow[T_0]{*} H_0 \iff \sigma(G_0) \xleftrightarrow[T_3]{*} \sigma(H_0). \tag{9}$$

This mapping $\sigma$ will be effective, so it will provide a reduction of the word problem for $T_0$ to the word problem for $T_3$. Taking for $T_0$ any Thue system with undecidable word problem we shall obtain a Thue system $T_3$ with 3 relation and undecidable word problem. If for the system $T_0$ there is no algorithm to decide, given a word $H$, whether it is equivalent to some *fixed* word $G$, then the problem of the equivalence to word $\sigma(G)$ is undecidable for system $T_3$.

In order to reach our ultimate goal, Thue system $T_3$, we shall construct two auxiliary Thue systems $T_1$ and $T_2$ in some alphabets $A_1$ and $A_2$ and for $i = 0, 1, 2$ define mappings

$$\tau_i : A_i^* \longrightarrow A_{i+1}^*$$

such that for any two words $G_0$ and $H_0$ from $A_0^*$

$$G_i \xleftrightarrow[T_i]{*} H_i \iff \tau_i(G_i) \xleftrightarrow[T_{i+1}]{*} \tau_i(H_i) \tag{10}$$

where

$$G_{i+1} \equiv \tau_i(G_i)$$

and

$$H_{i+1} \equiv \tau_i(H_i).$$

It is clear that we can define

$$\sigma(G_0) \equiv G_3$$

and obtain (9).

Each mapping $\tau_i$ will be defined via a word $Z_{i+1}$ from $A_{i+1}^*$ and an *encoding*

$$\rho_i : A_i \longrightarrow A_{i+1}^*.$$

Such an encoding can be extended in a natural way to a mapping

$$\rho_i : A_j^* \longrightarrow A_{i+1}^*,$$

namely, if

$$G_i \equiv g_1 \cdots g_k$$

then

$$\rho_i(G_i) \equiv \rho_i(g_1) \ldots \rho_i(g_k).$$

Mapping $\tau_i$ will be defined by

$$\tau_i(G_i) \equiv \rho_i(G_i) Z_{i+1}. \tag{11}$$

It is clear that to prove part $\Longrightarrow$ of the equivalence (10) it will be sufficient to prove that

$$G_i \xleftrightarrow[T_i]{} H_i \implies \tau_i(G_i) \xleftrightarrow[T_{i+1}]{*} \tau_i(H_i). \tag{12}$$

The proof of the inverse implications $\Longleftarrow$ in (10) will be based on defining mappings

$$\phi_i : A_{i+1}^* \to A_i^*$$

which will be semiinverse to $\tau_i$ in the sense that

$$\phi_i(\tau_i(G_i)) \equiv G_i.$$

These mappings $\phi_i$ will also provide mappings of derivations in $T_{i+1}$ into derivations in $T_i$ in the following sense: as soon as

$$\tau_i(G_i) \equiv W_0 \xleftrightarrow[T_{i+1}]{} W_1 \xleftrightarrow[T_{i+1}]{} \dots \xleftrightarrow[T_{i+1}]{} W_k \tag{13}$$

is a derivation in $T_{i+1}$,

$$G_i \equiv \phi_i(W_0) \xleftrightarrow[T_i]{} \phi_i(W_1) \xleftrightarrow[T_i]{} \dots \xleftrightarrow[T_i]{} \phi_i(W_k) \tag{14}$$

will be a derivation in $T_i$.

## 2.2 System $T_1$

Our first goal is the reduction of the word problem for $T_0$ to the word problem for a Thue system $T_1$ with defining relations

$$\begin{gathered} P_1 \longleftrightarrow Q_1 \\ \vdots \\ P_{t_1} \longleftrightarrow Q_{t_1} \end{gathered} \tag{15}$$

such that words $P_1, \dots, P_{t_1}$, and words $Q_1, \dots, Q_{t_1}$ have equal and non-zero lengths, i.e.,

$$|P_1| = \dots = |P_{t_1}| = p \neq 0$$

and

$$|Q_1| = \dots = |Q_{t_1}| = q \neq 0.$$

To this end we extend alphabet $A_0$ by a new genetator, namely, we define

$$A_1 = \{a_1, \dots, a_n, a_{n+1}\}.$$

Let

$$p = \max\{|R_1|, \dots, |R_{t_0}|, |S_1|, \dots, |S_{t_0}|\} + 1,$$
$$q = p + 1.$$

For each relation $R_i \longleftrightarrow Q_i$ of system $T_0$ we will include into (15) the relation

$$R_i a_{n+1}^{p-|R_i|} \longleftrightarrow S_i a_{n+1}^{q-|S_i|}. \tag{16}$$

Also for each generator $a_j$ from $A_1$ we include into (15) the relation

$$a_j a_{n+1}^{p-1} \longleftrightarrow a_{n+1} a_j a_{n+1}^{q-2} \tag{17}$$

Thanks to the latter relations, for every $j$ we have

$$a_j a_{n+1}^q \overset{*}{\underset{T_1}{\longleftrightarrow}} a_{n+1}^q a_j a_{n+1}^q$$

and by induction on the length of an arbitrary word $G$

$$G a_{n+1}^q \overset{*}{\underset{T_1}{\longleftrightarrow}} a_{n+1}^q G' a_{n+1}^q$$

for some word $G'$.

Encoding $\rho_0$ is the identity $\rho_0(a_j) \equiv a_j$, and mapping $\tau_0$ is defined by (11) with $Z_1 \equiv a_{n+1}^q$.

Let us check (12) for $i = 0$. Case $G_0 \equiv H_0$ is trivial. Let now $G_0 \equiv XR_iY$, $H_0 \equiv XS_iY$. We have:

$$\begin{aligned} \tau_0(G_0) &\equiv XR_iYa_{n+1}^q \\ &\overset{*}{\underset{T_1}{\longleftrightarrow}} XR_ia_{n+1}^qY'a_{n+1}^q \\ &\underset{T_1}{\longleftrightarrow} XS_ia_{n+1}^qY'a_{n+1}^q \\ &\overset{*}{\underset{T_1}{\longleftrightarrow}} XS_iYa_{n+1}^q \\ &\equiv \tau_0(H_0). \end{aligned}$$

Case $G_0 \equiv XS_iY$, $H_0 \equiv XR_iY$ follows by the symmetry of relation $\overset{*}{\underset{T_1}{\longleftrightarrow}}$.

The inverse mapping $\phi_0$ is nothing more than the projection of alphabet $A_1$ onto alphabet $A_0$, i.e., $\phi_0(G_1)$ is the result of deleting all occurences of letter $a_{n+1}$. To check that $\phi_0$ does transform a derivation (13) in $T_1$ into a derivation (14) in $T_0$, it suffies to note that

$$\phi_0(XP_iY) \equiv \phi_0(X)\phi_0(P_i)\phi_0(Y)$$

and,

$$\phi_0(XQ_iY) \equiv \phi_0(X)\phi_0(Q_i)\phi_0(Y),$$

and either $\phi_0(P_i) \longleftrightarrow \phi_0(Q_i)$ is a defining relation of $T_0$ or $\phi_0(P_i) \equiv \phi_0(Q_i)$, depending on whether $P_i \longleftrightarrow Q_i$ is a relation of type (16) or (17).

### 2.3 System $T_2'$

Let $A_2' = \{a, b\}$ and $\rho_1$ be the encoding such that

$$\rho_1 : a_j \mapsto aab^j ab^{n+2-j}. \tag{18}$$

Let $t$ be the smallest power of 2 which is not less than $t_1$,

$$t = 2^u.$$

Let $P_i$ and $Q_i$ denote words $P_{t_1}$ and $Q_{t_1}$ whenever $i \geq t_1$. Let for $i = 1, \ldots, t$

$$\begin{aligned} L_i &\equiv \rho_1(P_i), \\ M_i &\equiv \rho_1(Q_i). \end{aligned}$$

We first consider system $T_2'$ with $t$ relations

$$\begin{aligned} L_1 &\longleftrightarrow M_1, \\ &\vdots \\ L_t &\longleftrightarrow M_t. \end{aligned} \tag{19}$$

It is evident

$$G_1 \underset{T_1}{\overset{*}{\longleftrightarrow}} H_1 \iff \rho_1(G_1) \underset{T_2'}{\overset{*}{\longleftrightarrow}} \rho_1(H_1)$$

thanks to the fact that the $P$'s and $Q$'s are non-empty words. The role of our choice of $t$ as a power of 2 will become clear later.

A reduction of the word problem for arbitrary Thue system to the word problem for a Thue system in a two letter alphabet was done by Hall [15] using an encoding different from (18). Our special choice of $\rho_1$ pursued several goals. First, similar to system $T_1$ the lengths of left-hand sides and the right-hand sides of all relation in (19) are equal, respectively, to some numbers $l$ and $m$. Second, we shall use later the fact that a word of the form $\rho_1(G_1)$ does not contain more than two consecutive $a$'s, and all couples of consecutive $a$'s originate from the first two $a$'s in enconding (18).

### 2.4 System $T_2$

Our new system $T_1$ has more relations than the original system $T_0$. The next system $T_2$ will have only 5 relation independent of the number of relation in $T_0$. The idea is to compress all relation (19) into a single one. Let

$$\begin{aligned} L_i &\equiv l_{i1} \ldots l_{il}, \\ M_i &\equiv m_{i1} \ldots m_{im}, \end{aligned}$$

$$\begin{aligned} L &\equiv l_{11} l_{21} \ldots l_{t1} \ldots l_{1l} l_{2l} \ldots l_{tl}, \\ M &\equiv m_{11} m_{21} \ldots m_{t1} \ldots m_{1m} m_{2m} \ldots m_{tm}, \end{aligned}$$

and let $T_2$ be the Thue system in alphabet

$$A_2 = \{a, b, e\}$$

with relations

$$eaa \leftrightarrow ae, \tag{20}$$
$$eab \leftrightarrow be, \tag{21}$$
$$eba \leftrightarrow ae, \tag{22}$$
$$ebb \leftrightarrow be, \tag{23}$$
$$L \leftrightarrow M. \tag{24}$$

Thanks to relations (20)–(23)

$$e^u X x \overset{*}{\underset{T_2}{\longleftrightarrow}} x e^u$$

for any $x$ from $A'_2$ and any word $X$ from $A'^*_2$ having length $t-1$. By induction on the length of an arbitrary word $G$ we have

$$G e^u \overset{*}{\underset{T_2}{\longleftrightarrow}} e^u G'$$

for some word $G'$, in particular, for any $i$

$$L_i a e^u \overset{*}{\underset{T_2}{\longleftrightarrow}} e^u a^{t-i} L a^i.$$

Let $\tau_1$ be defined by (11) and (18) with $Z_2 \equiv ae^u$. Let us check (12) for $i=1$. For the nontrivial case $G_1 \equiv XP_iY$, $H_1 \equiv XQ_iY$ we have:

$$\begin{aligned}
\tau_1(G_1) &\equiv \rho_1(X) L_i \rho_1(Y) a e^u \\
&\overset{*}{\underset{T_2}{\longleftrightarrow}} \rho_1(X) L_i a e^u Y' \\
&\overset{*}{\underset{T_2}{\longleftrightarrow}} \rho_1(X) e^u a^{t-i} L a^i Y' \\
&\underset{T_2}{\longleftrightarrow} \rho_1(X) e^u a^{t-i} M a^i Y' \\
&\overset{*}{\underset{T_2}{\longleftrightarrow}} \rho_1(X) M_i a e^u Y' \\
&\overset{*}{\underset{T_2}{\longleftrightarrow}} \rho_1(X) M_i \rho_1(Y) a e^u \\
&\equiv \tau_1(H_1).
\end{aligned}$$

The definition of $\phi_1$ is the most tricky part of the whole construction. This mapping is defined by *semiThue system with priorities* having three blocks of rules:

$$\begin{aligned}
eaa &\longrightarrow ae, \\
eab &\longrightarrow be, \\
eba &\longrightarrow ae, \\
ebb &\longrightarrow be;
\end{aligned}$$

$$aab^1ab^{n+1} \longrightarrow a_1,$$

$$\vdots$$

$$aab^{n+1}ab \longrightarrow a_{n+1};$$

$$e \longrightarrow \Lambda,$$
$$a \longrightarrow \Lambda,$$
$$b \longrightarrow \Lambda.$$

The rules of the first block correspond to relations (20)–(23); the rules of the second block correspond to the encoding (18). To calculate $\phi_1(G)$ one has at first apply the rules from the first block until reaching a word $G'$ to which none of the rules from the first block can be applied any longer; after that rules of the second block should be applied as long as possible producing a word $G''$; finally, the rules from the third block should finish transforming $G$ into a word $G'''$ from $A_1^*$. It is not difficult to understand that each of these three blocks of rules constitutes a *confluent, strictly length-decreasing* semi Thue system so these words $G'$, $G''$ and $G'''$ are completely determined by word $G$ and do not depend on the particular way in which the rules were applied; these words will be denoted respectively by $\phi'(G)$, $\phi''(G)$ and $\phi_1(G)$.

Suppose that we have a derivation (13) with $i = 1$. Let us prove by induction on $k$ that in such a case

$$\rho_1(G_1)ae^u \equiv \phi'(W_0) \underset{T_2'}{\longleftrightarrow} \phi'(W_1) \underset{T_2'}{\longleftrightarrow} \dots \underset{T_2'}{\longleftrightarrow} \phi'(W_k) \tag{25}$$

is a derivation in $T_2'$ (we consider now $T_2'$ as a system in alphabet $A_2$ rather than in $A_2'$).

The base $k = 0$ is evident. Suppose now that (25) is a derivation, $W_k \equiv XUY$ and $U \longleftrightarrow V$ or $V \longleftrightarrow U$ is a relation of system $T_2$. If it is one of the four relations (20)–(23), then clearly $\phi'(XVY) \equiv \phi'(W_k)$. Without loss of generality it remains to consider only the case $U \equiv L$, $V \equiv M$. Being equivalent in $T_2'$ to $\rho_1(G_1)ae^u$, the word $\phi'(W_k)$ does not contain more than two consecutive $a$'s. On the other hand, word $L$ begins with $2t = 2^{u+1}$ $a$'s and hence word $X$ should contain at least $u$ occurences of letter $e$. It cannot contain more than $u$ occurences because word $W_k \equiv XUV$ is equivalent to $\tau_1(G_1)$ in $T_2$. The difference of the lengths of words $L$ and $M$ is divisible by $t$, so it is easy to understand that, for some words $X'$, $Y'$ and number $i$,

$$\phi'(W_k) \equiv \phi'(XLY) \equiv X'L_iY'e^u$$

and

$$\phi'(XMY) \equiv X'M_iY'e^u$$

so these two words are equivalent in $T_2'$. The induction is completed.

Now that we know that (25) is a derivation in $T_2'$, we see that (14) is a derivation in $T_1$.

## 2.5 System $T_3$

Let $A_3 = \{c, d\}$ and let $\rho_2$ be the following encoding:

$$\begin{aligned} \rho_2(a) &\equiv c, \\ \rho_2(b) &\equiv cd, \\ \rho_2(e) &\equiv dd. \end{aligned} \tag{26}$$

We define $\tau_2$ by (11) with $Z_3$ being the empty word. Let $T_3$ be the system

$$\begin{aligned} ddcc &\longleftrightarrow cdd, \\ ddcdc &\longleftrightarrow cdd, \\ \rho_2(L) &\longleftrightarrow \rho_2(M). \end{aligned}$$

To check (12) for $i = 2$ it suffices to observe that

$$\begin{aligned} \rho_2(eaa) &\underset{T_3}{\longleftrightarrow} \rho_2(ae), \\ \rho_2(eab) &\underset{T_3}{\longleftrightarrow} \rho_2(be), \\ \rho_2(eba) &\underset{T_3}{\longleftrightarrow} \rho_2(ae), \\ \rho_2(ebb) &\underset{T_3}{\longleftrightarrow} \rho_2(be). \end{aligned}$$

The inverse mapping $\phi_2$ can be defined as follows:

$$\begin{aligned} \phi_2(\Lambda) &\equiv \Lambda, \\ \phi_2(d) &\equiv \Lambda, \\ \phi_2(Gc) &\equiv \phi_2(G)a, \\ \phi_2(Gcd) &\equiv \phi_2(G)b, \\ \phi_2(Gdd) &\equiv \phi_2(G)e. \end{aligned}$$

Besides the required identity

$$\phi_2(\tau_2(G_2)) \equiv G_2$$

for every word $G$ from $A_3^*$ we have either

$$\tau_2(\phi_2(G)) \equiv G$$

or

$$d\tau_2(\phi_2(G)) \equiv G.$$

Suppose that we have a derivation (13) with $i = 2$. Let us prove by induction on $k$ that in such a case (14) is a derivation in $T_2$. The base of induction, $k = 0$, is trivial. Let $W_k \equiv XUY$ where either $U \longleftrightarrow V$ or $V \longleftrightarrow U$ is a defining relation of $T_3$. Let

$$y\tau_2(\phi_2(Y)) \equiv Y$$

where either $y \equiv \Lambda$ or $y \equiv d$. We have:

$$\phi_2(W_k) \equiv \phi_2(XUY) \equiv \phi_2(X)\phi_2(Uy)\phi_2(Y)$$

and

$$\phi_2(XVY) \equiv \phi_2(X)\phi_2(Vy)\phi_2(Y),$$

and either $\phi_2(Uy) \longleftrightarrow \phi(Vy)$ or $\phi_2(Vy) \longleftrightarrow \phi(Uy)$ is a defining relation of $T_2$ with the exception of the case when $y \equiv d$ and $U \longleftrightarrow V$ or $V \longleftrightarrow U$ is the "long" relations $\rho_2(L) \longleftrightarrow \rho_2(M)$. But this case is impossible because, as it was shown above, the word $\phi_2(X)$ should contain all $u$ occurences of letter $e$ while the word $\phi_2(Uy)$ ends with $e$.

## References

1. S. I. Adian. On P. S. Novikov's and his disciple's investigations on algorithmical problems in algebra (in Russian). *Trudy Mat. Inst. Steklov.*, 133:23–32, 1973.
2. S. I. Adian and G. S. Makanin. Investigations on algorithmical problems in algebra (in Russian). *Trudy Mat. Inst. Steklov.*, 168:197–217, 1984.
3. W. W. Boone. The word problem. *Ann. Math.*, 70(2):207–265, 1959.
4. W. W. Boone, D. Collins, and Ju. V. Matijasevič. Embedding into semigroups with only a few defining relations. In J. E. Fenstad, editor, *Proceedings of the Second Scandinavian Logic Symposium*, volume 63 of *Studies in Logic and the Foundations of Mathematics*, pages 27–40, Amsterdam, 1971. North-Holland.
5. W. W. Boone and D. J. Collins. Embeddings into groups with only a few defining relations. *J. Austral. Math. Soc.*, 18(1):1–7, 1974.
6. V. V. Borisov. Simple examples of groups with undecidable word problem (in Russian). *Mat. Zametki*, 6(5):521–532, 1969.
7. J. L. Britton. The word problem. *Ann. Math.*, 77(1):16–32, 1963.
8. V. Claus. Die Grenze zwischen Entscheidbarkeit und Nichtentscheidbarkeit. *Fernstndeinkurs für die Fernuniversität Hagen*, Open University, Hagen, 1979.
9. V. Claus. Some remarks on PCP($k$) and related problems. *Bull. EATCS*, 12:54–61, 1980.
10. D. J. Collins. A simple presentation of a group with unsolvable word problem. *Illinois J. Math.*, 30(2):230–234, 1986.
11. D. J. Collins. Word and conjugacy problems in groups with only a few defining relations. *Z. Math. Logik Grundlag. Math.*, 15(4):305–323, 1969.
12. M. Dehn. Über die Topologie des dreidimensionalen Raumes. *Math. Ann.*, 69:137–168, 1910.
13. M. Dehn. Über unendliche diskontinuierliche Gruppen. *Math. Ann.*, 71:116–144, 1912.
14. A. Ehrenfeucht and G. Rozenberg. On the (generalized) Post correspondence problem with lists of length 2. *Lect. Notes Comput. Sci.*, 115:408–416, 1981.
15. M. J. Hall. The word problem for semigroups with two generators. *J. Symbolic Logic*, 14:115–118, 1949.
16. G. Higman. Subgroups of finitely presented groups. *Proc. Roy. Soc.*, 262(1311):455–475, 1961.
17. G. Lallement. Presentations de monoides et problems algorithmiques. *Lecture Notes in Mathematics*, 586:136–144, 1977.
18. G. Lallement. The word problem for Thue rewriting systems. *This volume*, 1994.
19. J. Lockhart. Triviality problem for semigroups of deficiency 1. *J. Symbolic Logic*, 42(3):457–458, 1977.

20. W. Magnus. Das Identitäts problem für Gruppen mit einer definierenden Relation. *Math. Ann.*, 106:295–307, 1932.

21. G. S. Makanin. On the word problem for finitely presented semigroups. *Soviet Math. Doklady*, 7:1478–1480, 1966.

22. C. Marché. On ground AC-completion. *Lect. Notes Comput. Sci.*, 488:411–422, 1991.

23. A. A. Markoff. Impossibility of certain algorithms in the theory of associative systems (in Russian). *Dokl. Akad. Nauk SSSR*, 55(7):587–590, 1947.

24. A. A. Markoff. Impossibility of certain algorothms in the theory of associative systems (in Russian). *Dokl. Akad. Nauk SSSR*, 58(3):353–356, 1947.

25. A. A. Markoff. Theory of algorithms (in Russian). *Trudy Mat. Inst. Steklov.*, 42, 1954. Translated in Israel Program of scientific Translations. Jerusalem, 1961. MR 17, 1038, MR 24, A2527.

26. Yu. Matiyasevich. Simple examples of undecidable associative calculi (in Russian). *Dokl. Akad. Nauk SSSR*, 173:1264–1266, 1967. English translation in: *Soviet Math. Dokl.*, 8:555–557, 1967.

27. Yu. Matiyasevich. Simple examples of undecidable canonical calculi (in Russian). *Trudy Mat. Inst. Steklov.*, 93:50–88, 1967. English translation in: *Proc Steklov Inst. Math.*, 93:227–252, 1968.

28. V. L. Murskii. Unrecognizable properties of finite systems of identity relations (in Russian). *Dokl. Akad. Nauk SSSR*, 196(3):520–522, 1971.

29. P. S. Novikov. On algorithmical undecidability of the word problem in the theory of groups (in Russian). *Dokl. Akad. Nauk SSSR*, 85(4):709–712, 1952.

30. P. S. Novikov. On algorithmical undecidability of the word problem in the theory of groups (in Russian). *Trudy Mat. Inst. Steklov.*, 44, 1955.

31. J. J. Pansiot. A note on Post's correspondence problem. 12(5):233, 1981.

32. E. L. Post. A variant of recursively unsolvable problem. *Bull. Amer. Math. Soc.*, 52:264–268, 1946.

33. E. L. Post. Recursive unsolvability of a problem of Thue. *J. Symbolic Logic*, 12:1–11, 1947.

34. L. Priese. Über ein 2-dimensionales Thue-System mit zwei Regeln und unentscheidbaren Wortproblem. *Z. Math. Logik Grundlag. Math.*, 25(2):179–192, 1979.

35. D. Scott. A short recursively unsolvable problem. *J. Symbolic Logic*, 21:111–112, 1956.

36. G. Sénizergues. Some undecidable termination problem for semi-Thue systems. To appear in *Lect. Notes Comput. Sci.*, 1994

37. A. Thue. Probleme über Veränderungen von Zeichenreihen nach gegebenen Regeln. *Skrifter utgit av Videnskapsselskapet i Kristiania*, I. Matematisk-naturvidenskabelig klasse, 10, 34pp., 1914. Reprinted in: A. Thue. Selected Mathematical Papers. Oslo, 1977, 493–524.

38. G. S. Tseitin. An associative calculus with undecidable problem of equivalence (in Russian). *Dokl. Akad. Nauk SSSR*, 107(3):370–371, 1956.

39. G. S. Tseitin. An associative calculus with undecidable problem of equivalence (in Russian). *Trudy Mat. Inst. Steklov.*, 52:172–189, 1958.

40. M. K. Valiev. Universal group with twenty-one defining relation. *Discrete Math.*, 17:207–213, 1977.

41. M. K. Valiev. Examples of universal finitely-presented groups (in Russian). *Dokl. Akad. Nauk SSSR*, 211(2):265–268, 1973.

42. M. K. Valiev. On polynomial reducibility of word problem under embedding of recursively presented groups in finitely presented groups. In J. Bečvář, editor, *Mathematical Foundations of Computer Science 1975*, volume 32 of *Lect. Notes Comput. Sci.*, pages 432–438. Springer-Verlag, 1975.

43. A. Yasuhara. The solvability of the word problem for certain semigroups. *Proc. Amer. Math. Soc.*, 26(4):645–650, 1970.