DECISION PROBLEMS FOR SEMI-THUE SYSTEMS WITH A FEW RULES*

Yuri Matiyasevich (Steklov institute of Mathematics, St Petersburg)†
and Géraud Sénizergues (LaBRI , Bordeaux)

Abstract

For several decision problems about semi-Thue systems , we try to locate the frontier between the decidable and the undecidable from the point of view of the number of rules. We show that the the Termination Problem, the U-Termination Problem, the Accessibility Problem and the Common-Descendant Problem are undecidable for 3 rules STS . As a corollary we obtain the undecidability of the Post-Correspondence Problem for 7 rules.

Keywords: Decidability ; bounds; semi-Thue system; termination; accessibility; Post Correspondence Problem.

1 Introduction

It is known that the *termination* problem as well as the *uniform-termination* problem are undecidable for finite semi-Thue systems ([10, Theorem 2.2 p.70], [10, §2 p.88-93], [14, point (8) p.227], [16]), and it follows that the same problems are undecidable for more general types of rewriting systems (e.g. for the general word rewriting systems investigated in [22], for term rewriting systems, for graph rewriting systems, ...). These general undecidability results have been refined in several directions :

- there exists a *word* rewriting-system, with only *one rule* and undecidable termination-problem ([22, p.98]) (this system is not semi-Thue, in particular it is not linear and uses more than two variables).
- there exists a *term* rewriting-system, with only *one rule* and undecidable termination-problem ([9], improving upon the previous bound 2 given in [11]).

[1]An extended abstract of this work appeared as [25]

[2]This work has been partly realized while the first author was visiting the LaBRI, with financial support of the CNRS.

- the uniform-termination problem for *length-preserving* semi-Thue systems is still undecidable [6, 29]; this result implies the undecidability of the uniform-termination problem for the finite graph rewriting systems studied in [2], and also the undecidability of *simple* termination for finite semi-Thue systems([27]).

As well it is known that the *accessibility* problem is undecidable for finite semi-Thue systems. In particular the *word* problem for a semi-Thue system S (which is nothing else than the accessibility problem for the symetrized system $S \cup S^{-1}$), is undecidable for finite semi-Thue systems. This problem is usually called the *word*-problem for finitely presented monoids.

Many works have been devoted to the refinement of this undecidability result and its analogue for finitely presented groups:

- there exists a Thue-system, with only *3 rules* (here the rules are unoriented) and undecidable word problem ([21, 4, 24]).
- there exists a group presentation, with only *12 rules* and undecidable word-problem ([5, 8]).

On the other hand many works have been devoted to the discovery of algorithms solving the word problem for subclasses of semi-Thue systems (see [1, 23, 19] for detailed historical accounts and a wide bibliography).

We establish here *new tight upper bounds* on the number of rules of semi-Thue systems for which these problems might be decidable:

1. there exists a semi-Thue system having 3 rules and with undecidable termination problem (Theorem 4.1), thus improving the bound 10 of [29],
2. the uniform-termination problem is undecidable for finite semi-Thue systems having 3 rules (Theorem 4.2), thus improving the bound 9 of [29],
3. there exists a semi-Thue system having 3 rules and with undecidable accessibility problem (Theorem 4.3), thus improving the bound 5 resulting from [21] and [28]. This system has also undecidable common descendant problem.

Using a reduction constructed in [7], we then prove that the Post-correspondence problem is undecidable for 7 pairs of words, thus improving the bound 9 of [28], (we recall that the PCP is decidable for 2 pairs of words ([13])).

Up to our knowledge, it is still unknown whether every 1 rule semi-Thue system has a decidable termination-problem : this question, which has been part of the "folklore" for some years, is formaly raised in [9, p.410] and partial answers are given in [18, 26, 32, 30, 31]. Hence we have reduced the "area of uncertainty" for this problem to the interval $[1, 2]$ (this is true as well for the u-termination problem). The accessibility problem is obviously decidable for the 1 rule semi-Thue systems. Hence we have reduced the "area of uncertainty" for this problem just to the singleton $\{2\}$. Let us notice that the acessibility-problem for 2 rules STS contains as a sub-problem the word problem for 1 rule Thue systems (i.e. the rule is unoriented), which is an important open problem in the theory of semi-groups (see [19, part II] for a survey).

At last, let us notice that the proofs of our results show the existence of a *universal* semi-Thue system (in a suitable precise sense) which has only 3 rules. In this form , our work is related to all the efforts towards defining as small as possible universal devices (see e.g. [20]).

2 Preliminaries

2.1 Vocabulary, Notation

2.1.1 Words

Given an alphabet X, by $(X^*, ., \epsilon)$ we denote the *free monoid* generated by X (where . is the concatenation-product and ϵ is the empty word). As X^* is embedded in $F(X)$ (the *free group* generated by X) we sometimes use the notation u^{-1} for the *inverse* in $F(X)$ of a word u.

2.1.2 Semi-Thue systems

A *semi-Thue* system over X is a subset $S \subseteq X^* \times X^*$. By $\longrightarrow_S$ we denote the binary relation $\forall f, g \in X^*, f \longrightarrow_S g$ iff there exists $(u, v) \in S$ and $\alpha, \beta \in X^*$ such that $f = \alpha u \beta, g = \alpha v \beta$. $\longrightarrow_S$ is the *one-step* rewriting relation generated by S. The pair (α, u) is called the *redex* used in the rewriting-step $f \longrightarrow_S g$.

$\xrightarrow{i}_S$ (where $i \in \mathbb{N}$), $\xrightarrow{*}_S$, $\xrightarrow{+}_S$ are then defined as usual from $\longrightarrow_S$ (see [15]). By $\longleftrightarrow_S$ we denote the binary relation :$\forall f, g \in X^*, f \longleftrightarrow_S g$ iff $f \longrightarrow_S g$ or $g \longrightarrow_S f$. $\xleftrightarrow{i}_S$, $\xleftrightarrow{*}_S$, $\xleftrightarrow{+}_S$ are then defined as usual from $\longleftrightarrow_S$. A *derivation* (mod S) is a sequence $D = (w_i, \alpha_i, u_i, v_i)_{i \in I}$ where I is a non-empty beginning section of $\mathbb{N}$ (i.e. $I = [0, n]$ for some $n \in \mathbb{N}$ or $I = \mathbb{N}$), for every $i \in I, w_i \in X^*$, $(u_i, v_i) \in S$ and $(i + 1 \in I) \implies (\alpha_i u_i$ is a prefix of $w_i, w_{i+1} = \alpha_i v_i (\alpha_i u_i)^{-1} w_i)$. The *length* of D is n (if $I = [0, n]$) or ∞ (if $I = \mathbb{N}$). D is said to be a *rl-derivation* (rl stands for right-to-left) iff, $\forall i \in I - \{0\}, |\alpha_i| < |\alpha_{i-1} v_{i-1}|$. In order to abbreviate the notation we often (incorrectly) drop the data α_i, u_i, v_i in the definition of a derivation D.
We remind the reader that a semi-Thue system $S \subseteq X^* \times X^*$ is said *confluent* iff:

$$\forall u, v, w \in X^*, \text{ if } u \xrightarrow{*}_S v \text{ and } u \xrightarrow{*}_S w \text{ then}$$

$$\exists u' \in X^* \text{ such that } v \xrightarrow{*}_S u' \text{ and } w \xrightarrow{*}_S u'.$$

When S is both confluent and u-terminating, we call *total reduction* associated to S the map ρ_S sending every word in X^* on its unique normal form :

$$\forall w \in X^*, w \xrightarrow{*}_S \rho_S(w) \text{ and } w \text{ is irreducible} \pmod{S}.$$

2.1.3 Algorithmic problems

The following algorithmic problems on semi-Thue systems are classical.

(TP)*termination problem* **:** the termination-problem for the alphabet X and the semi-Thue system $S \subseteq X^* \times X^*$ is the following

instance: $w \in X^*$

question : does every derivation (mod S) starting on w have finite length ? (when the answer is "yes", we say that S *terminates* on w).

(UTP)*uniform termination problem* **:** the uniform-termination problem for a class $\mathcal{S}$ of semi-Thue systems is the following

instance: an alphabet X and a finite semi-Thue system $S \subset X^* \times X^*$ which belongs to $\mathcal{S}$.

question : does every derivation (mod S) starting from a word in X^* have finite length ?
(when the answer is "yes" we say that S is *uniformly-terminating*, sometimes abbreviated as u-terminating).

(ACP)*accessibility problem* **:** the accessibility-problem for the alphabet X and the semi-Thue system $S \subseteq X^* \times X^*$ is the following

instance : two words $w_1, w_2 \in X^*$

question : $w_1 \xrightarrow{*}_S w_2$?

(WP)*word-problem* **:** the word problem for the alphabet X and the semi-Thue system $S \subseteq X^* \times X^*$ is the following

instance: two words $w_1, w_2 \in X^*$

question : $w_1 \xleftrightarrow{*}_S w_2$?

(IAP)*individual accessibility problem* **:** the individual accessibility-problem for the alphabet X, the semi-Thue system $S \subset X^* \times X^*$ and the word $w_0 \in X^*$ is the following

instance : one word $w \in X^*$

question : $w \xrightarrow{*}_S w_0$?

(IWP) *individual word-problem* **:** the individual word problem for the alphabet X , the semi-Thue system $S \subseteq X^* \times X^*$ and the word $w_0 \in X^*$ is the following

instance: one word $w \in X^*$

question : $w \xleftrightarrow{*}_S w_0$?

For more information the reader should refer to [17, 3] (about semi-Thue systems) or [11, 12] (about termination problems).

2.2 Some useful results and systems

2.2.1 Undecidability results

It is known that the *uniform-termination* problem for finite semi-Thue systems is undecidable ([14, point (8) p.227] or [16]). More recently it has been established that the uniform-termination problem for *length-preserving* semi-Thue systems is undecidable too ([6] or [29]). More precisely, the proof of [14] shows that the uniform-termination problem is not co-semi-decidable while the proofs of [6, 29] show that the uniform termination problem is not semi-decidable, even for the class of length-preserving semi-Thue systems. (Of course, the uniform termination problem is co-semi-decidable for the class of length-preserving semi-Thue systems).

It is also well-known that there exists some fixed finite semi-Thue system with undecidable *termination*-problem (this follows from the undecidability of the "*halting*-problem" for some fixed Turing machine [10, Theorem 2.2 p.70] and the translation of Turing machines into semi-Thue systems [10, §2 p.88-93] or [16]).

The following reduction of the IWP to the PCP will be useful.

Theorem 2.1 ([7]) *The accessibility problem for semi-Thue systems with n rules reduces to the Post correspondence problem for $n+4$ pairs of words.*

2.2.2 System $\mathcal{T}$

The existence of some fixed semi-Thue system with undecidable termination problem is well-known (see §2.2.1). Let us give below a rather small system $\mathcal{T}$ which has undecidable termination problem.(The *proof* of the fact that $\mathcal{T}$ has the announced property is not given here).
Let us define alphabets $A = \{a_1, a_2\}$,$\bar{A} = \{\bar{a_1}, \bar{a_2}\}$, $B = \{b_1, b_2\}$.
$\mathcal{T}$ is the system over the alphabet $A \cup \bar{A} \cup B \cup \{c, d, e\}$ consisting of the rules :

$$\begin{array}{lll} a_i c \longrightarrow c a_i, & a_i b_j \longrightarrow b_j a_i, & c c a_i b_i \longrightarrow c c b_i d, \\ d a_i b_i \longrightarrow b_i d, & dc \longrightarrow ce, & e b_i \longrightarrow \bar{a}_i b_i e, \\ e c c a_i \longrightarrow \bar{a}_i e c c, & ecc \longrightarrow cc, & b_j \bar{a}_i \longrightarrow \bar{a}_i b_j, \\ c \bar{a}_i \longrightarrow \bar{a}_i c, & \bar{a}_i \longrightarrow a_i. & \end{array}$$

2.2.3 System $\mathcal{C}$

Ceitin has constructed the following small semi-Thue system $\mathcal{C}$ which has undecidable individual word-problem . $\mathcal{C}$ is the system over the alphabet $\{a, b, c, d, e\}$ consisting of the rules:

$$ac \longleftrightarrow ca, \qquad ad \longleftrightarrow da, \qquad bc \longleftrightarrow cb, \qquad bd \longleftrightarrow db,$$
$$eca \longleftrightarrow ce, \qquad edb \longleftrightarrow de,$$
$$cdca \longleftrightarrow cdcae, \qquad caaa \longleftrightarrow aaa, \qquad daaa \longleftrightarrow aaa$$

Ceitin has shown that the IWP for $(\mathcal{C}, aaa)$ is undecidable.

2.2.4 rl-derivations

Lemma 2.2 ([29]) *Let S be a finite subset of $X^+ \times X^*$ and let $D = (w_i, \alpha_i, u_i, v_i)_{i \in I}$ be a derivation (mod S).*

1. *If I is finite, there exists some permutation $\sigma : I \rightarrow I$ and some sequence $(w'_i, \alpha'_i)_{i \in I}$ such that, $w'_0 = w_0$ and $(w'_i, \alpha'_i, u_{\sigma(i)}, v_{\sigma(i)})_{i \in I}$ is a rl-derivation (mod S)*
2. *If $I = N$, there exists some injection $\sigma : I \rightarrow I$ and some sequence $(w'_i, \alpha'_i)_{i \in I}$ such that, $w'_0 = w_0$ and $(w'_i, \alpha'_i, u_{\sigma(i)}, v_{\sigma(i)})_{i \in I}$ is a rl-derivation (mod S).*

3 Constructions

3.1 Definitions

We start with some given semi-Thue system S_0 over a finite alphabet A_0, consisting of rules:

$$\begin{array}{ccc} \overline{u_1} & \longrightarrow & \overline{v_1} \\ & \vdots & \\ \overline{u_{n_0}} & \longrightarrow & \overline{v_{n_0}} \end{array}$$

We then construct

- a sequence of alphabets A_i and semi-Thue systems S_i over A_i (for $1 \leq i \leq 5$)
- maps $\tau_i : {A_i}^* \rightarrow {A_{i+1}}^*$, (for $0 \leq i \leq 4$)
- partial maps $\phi_i : {A_{i+1}}^* \rightarrow {A_i}^*$, (for $0 \leq i \leq 4$)

such that ϕ_i is always a left-inverse of τ_i in the sense that:

$$\forall w \in {A_i}^*, \phi_i(\tau_i(w)) = w.$$

Every τ_i is then an *encoding* (i.e. an injective map) while ϕ_i will be called a *decoding*. For technical reasons we need to introduce auxiliary homomorphisms $\eta, \delta, \gamma, \varphi, \pi, \psi$, partial maps $\bar{\varphi}, \bar{\psi}$ and semi-Thue systems $\bar{S}_i$ on the alphabet A_i (for $i \in \{3, 4\}$). The situation is summarized on figure 1.

The aim of subsection 3.2 will be then to prove statements of the two following general forms:

- for every words w, w' in A_i^*,

$$w \longrightarrow_{S_i} w' \Rightarrow \tau_i(w) \overset{*}{\longrightarrow}_{S_{i+1}} \tau_i(w')$$

- for every words w, w' in A_{i+1}^* (and fulfilling some suitable condition)

$$w \longrightarrow_{S_{i+1}} w' \Rightarrow \phi_i(w) \overset{*}{\longrightarrow}_{S_i} \phi_i(w').$$

3.1.1 System S_1

Let $X = A_0 \cup \{e\}$. Let us note $X = \{x_1, x_2, \ldots, x_h\}$ where $h = |X|$ and $x_h = e$. Let $l_)$ (resp. m_0) be the maximum length of the lefthand (resp. righthand) sides of the rules of S_0.
We define a new system T over X consisting of the rules:

$$\begin{array}{rcll} e^{l_0} & \longrightarrow & \epsilon & \\ \overline{u_1}\beta & \longrightarrow & \overline{v_1}\beta & (\text{ for all } \beta \in X^{l_0 - |\overline{u_1}|}) \\ & \vdots & & \\ \overline{u_{n_0}}\beta & \longrightarrow & \overline{v_{n_0}}\beta & (\text{ for all } \beta \in X^{l_0 - |\overline{u_{n_0}}|}). \end{array} \tag{1}$$

Let $A_1 = \{p, q\}$. Let $\eta : A_0^* \rightarrow A_1^*$ be the monoid homomorphism defined by

$$\forall i \in [1, n_0], \eta(x_i) = ppq^i pq^{h+1-i}.$$

The map $\tau_0 : {A_0}^* \rightarrow {A_1}^*$ is then defined by

$$\forall w \in {A_0}^*, \tau_0(w) = \eta(we^l)$$

The system S_1 consists of the rules

$$\begin{array}{rcll} \eta(e^{l_0}) & \longrightarrow & \epsilon & \\ \eta(\overline{u_1}\beta) & \longrightarrow & \eta(\overline{v_1}\beta) & (\text{ for all } \beta \in X^{l_0 - |\overline{u_1}|}) \\ & \vdots & & \\ \eta(\overline{u_{n_0}}\beta) & \longrightarrow & \eta(\overline{v_{n_0}}\beta) & (\text{ for all } \beta \in X^{l_0 - |\overline{u_{n_0}}|}). \end{array} \tag{2}$$

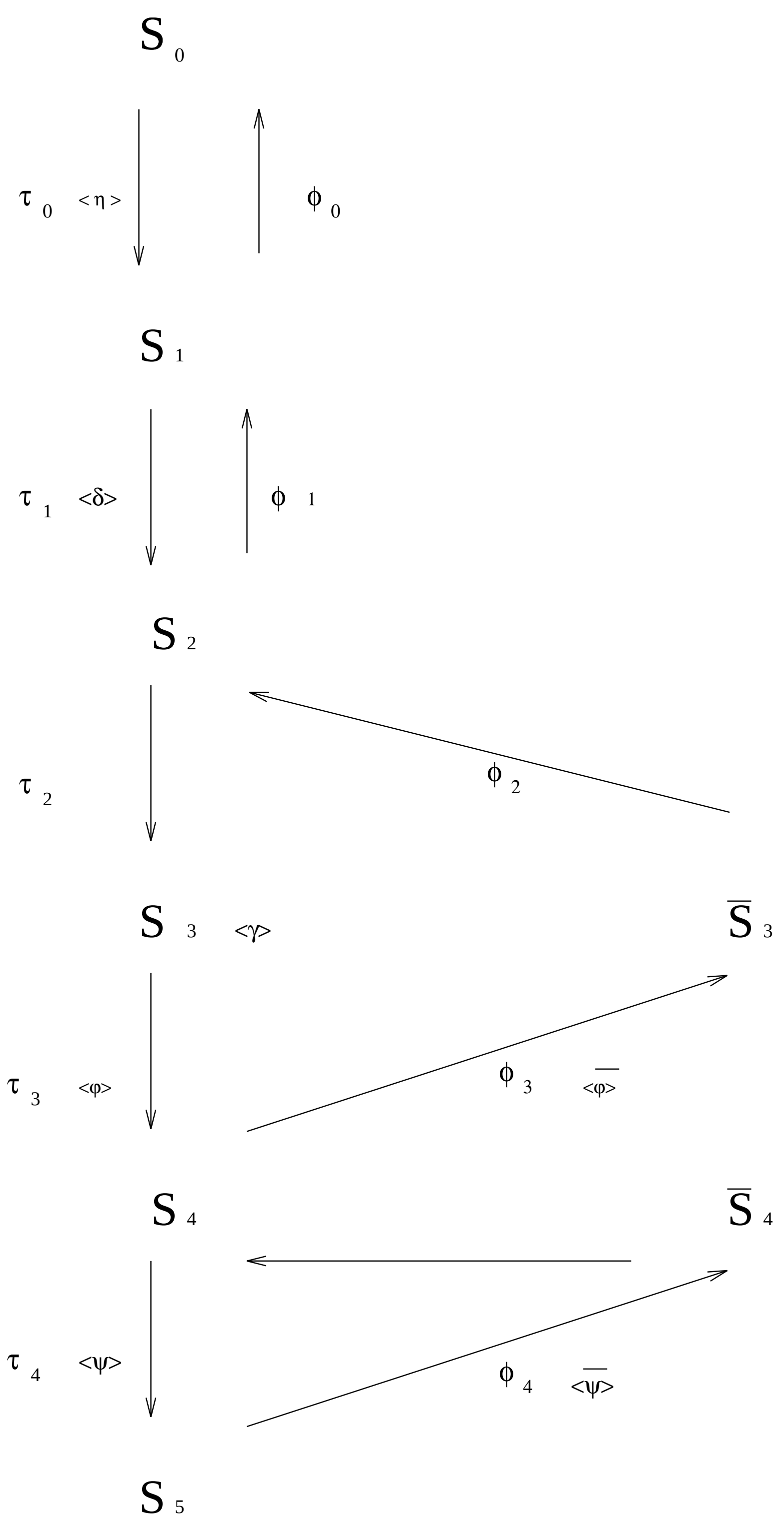

Figure 1: Systems, codings and decodings

The partial map $\phi_0 : {A_1}^* \to {A_0}^*$ is defined by means of the following two-level semi-Thue system $(T_0, {T'}_0)$ over $A_0 \cup \{e\} \cup A_1$:

$$
\begin{array}{rl}
T_0 : & ppq^i pq^{h+1-i} \longrightarrow x_i \quad (\text{ for } 1 \leq i \leq h) \\
{T'}_0 : & p \longrightarrow \epsilon \\
 & q \longrightarrow \epsilon \\
 & e \longrightarrow \epsilon.
\end{array} \tag{3}
$$

One can check that both semi-Thue systems $(T_0, {T'}_0)$ are noetherian and confluent. For every w in A_1^* we set:

$$\phi_0(w) = \rho_{{T'}_0}(\rho_{T_0}(w))$$

3.1.2 System S_2

The system S_1 defined above is now rewritten:

$$
\begin{array}{ccc}
u_0 & \longrightarrow & v_0 \\
 & \vdots & \\
u_n & \longrightarrow & v_n
\end{array}
$$

with

$$n \geq 0, \mu \geq 0, \nu \geq 0, |u_1| = \ldots = |u_n| = \mu; \quad |v_1| \leq \nu, \ldots, |v_n| \leq \nu$$

(where actually $\mu = l_0.(h+4), \nu = (l_0 + m_0).(h+4)$).
Let $A_2 = \{a, b\}$ and let $\delta : {A_1}^* \to {A_2}^*$ be the monoid homomorphism defined by

$$\delta(p) = b^2 abaa; \quad \delta(q) = b^2 aaba.$$

The map $\tau_1 : {A_1}^* \to {A_2}^*$ is then defined by

$$\forall w \in {A_1}^*, \tau_1(w) = \delta(w) b^2$$

The system S_2 consists of the rules

$$
\begin{array}{ccc}
\tau_1(u_0) & \longrightarrow & \tau_1(v_0) \\
 & \vdots & \\
\tau_1(u_n) & \longrightarrow & \tau_1(v_n) \\
b & \longrightarrow & a.
\end{array} \tag{4}
$$

We shall denote by $\mathcal{B}$ the semi-Thue system consisting of the single rule:

$$b \longrightarrow a.$$

The partial map $\phi_1 : {A_2}^* \rightarrow {A_1}^*$ is defined by means of the following two-level semi-Thue system $(T_1, {T'}_1)$:

$$\begin{array}{ll} T_1 : & b^2abaa \longrightarrow p \\ & b^2aaba \longrightarrow q \\ {T'}_1 : & a \longrightarrow \epsilon \\ & b \longrightarrow \epsilon. \end{array} \tag{5}$$

One can check that both semi-Thue systems (T_1, T_2) are noetherian and confluent. For every w in A_2^* we set:

$$\phi_1(w) = \rho_{{T'}_1}(\rho_{T_1}(w)).$$

3.1.3 System S_3

Let $A_3 = \{a, b, c, d\}$. The map $\tau_2 : {A_2}^* \rightarrow {A_3}^*$ is then defined by

$$\forall w \in {A_2}^*, \tau_2(w) = w$$

(τ_2 is simply the natural embedding of ${A_2}^*$ into ${A_3}^*$).
Let $\gamma : {A_2}^* \rightarrow {A_3}^*$ be the monoid homomorphism defined by:

$$\forall x \in A_2, \gamma(x) = cx.$$

The system S_3 consists of the rules

$$\begin{array}{rcl} \tau_1(u_0) & \longrightarrow & \gamma(\tau_1(v_0)c^{6(\nu-|v_0|)})c, \\ & \vdots & \\ \tau_1(u_n) & \longrightarrow & \gamma(\tau_1(v_n)c^{6(\nu-|v_n|)})c, \\ b & \longrightarrow & a, \\ c & \longrightarrow & \epsilon. \end{array} \tag{6}$$

$\bar{S}_3$ is the union of S_3 with the additional rule: $a \longrightarrow d$.

The partial map $\phi_2 : {A_3}^* \rightarrow {A_2}^*$ is the homomorphism from ${A_3}^*$ onto ${A_2}^*$ such that:

$$\phi_2(a) = a, \phi_2(b) = b, \phi_2(c) = \epsilon, \phi_2(d) = a.$$

3.1.4 System S_4

Let $A_4 = \{a, b, c, d, \hat{a}, \hat{b}, \hat{c}, \check{c}\}$.

Let $\varphi : {A_3}^* \rightarrow {A_4}^*$ be the unique monoid homomorphism such that:

$$\varphi(a) = \hat{a}b^n, \;\; \varphi(b) = \hat{b}b^n, \;\; \varphi(c) = \check{c}b^n, \;\; \varphi(d) = db^n.$$

The map $\tau_3 : A_3{}^* \to A_4{}^*$ is then defined by,

$$\forall w \in A_3{}^*, \tau_3(w) = \varphi(aw)c.$$

The system S_4 is defined as follows. Let us denote by $L_i \longrightarrow M_i$, for $0 \leq i \leq n$, the $n+1$ rules of system S_3 , by l_{ij} the $j+1$-th letter of L_i and by m_{ik} the $k+1$-th letter of M_i (for $0 \leq i \leq n, 0 \leq j \leq l, 0 \leq k \leq m$):

$$L_i = l_{i0} \dots l_{il}, \tag{7}$$
$$M_i = m_{i0} \dots m_{im}, \tag{8}$$

(where $l = 6.\mu - 1, m = 12.\nu$).
We then define the words

$$L = l_{00}l_{10} \dots l_{n0} \dots l_{0l}l_{1l} \dots l_{nl}, \tag{9}$$
$$M = m_{00}m_{10} \dots m_{n0} \dots m_{0m}m_{1m} \dots m_{nm}, \tag{10}$$

The system S_4 consists of the rules

$$\begin{array}{ccc}
 & L \longrightarrow M, & \\
\hat{b} \longrightarrow b, & \hat{b} \longrightarrow \hat{a}, & b \longrightarrow a, \\
\hat{a} \longrightarrow a, & a \longrightarrow d, & \check{c} \longrightarrow c, \\
c \longrightarrow d, & & \\
\hat{a}d^n cd^n a \longrightarrow \hat{a}b^n \hat{a}, & \hat{a}d^n cd^n b \longrightarrow \hat{a}b^n \hat{b}, & \hat{a}d^n cd^n c \longrightarrow \hat{a}b^n \check{c}, \\
\hat{b}d^n cd^n a \longrightarrow \hat{b}b^n \hat{a}, & \hat{b}d^n cd^n b \longrightarrow \hat{b}b^n \hat{b}, & \hat{b}d^n cd^n c \longrightarrow \hat{b}b^n \check{c}.
\end{array} \tag{11}$$

$\bar{S}_4$ is the union of S_4 with the additional rules:

$$\begin{array}{ccc}
\hat{c} \longrightarrow \hat{b} & \hat{c} \longrightarrow \check{c}, & \\
\hat{c}d^n cd^n a \longrightarrow \hat{c}b^n \hat{a}, & \hat{c}d^n cd^n b \longrightarrow \hat{c}b^n \hat{b}, & \hat{c}d^n cd^n c \longrightarrow \hat{c}b^n \check{c}.
\end{array} \tag{12}$$

Let $\pi : A_4{}^* \to A_3{}^*$ be the homomorphism defined by :

$$\pi(a) = a, \pi(\hat{a}) = a, \pi(b) = b, \pi(\hat{b}) = b, \pi(c) = c, \pi(\check{c}) = c, \pi(\hat{c}) = c, \pi(d) = d.$$

We introduce two subsets $Q_n, R_n \subseteq A_4^*$ by:

$$Q_n = \{w \in A_4{}^* \mid \forall i \in [1, |w|], w[i] \in \{\hat{a}, \hat{b}, \check{c}, \hat{c}\} \Rightarrow i \equiv 1 \pmod{n+1}\}$$

$$R_n = Q_n \cap \{w \in A_4{}^* \mid w[1] \in \{a, \hat{a}\}, |w| \equiv 1 \pmod{n+1} \text{ and } w[|w|] \in \{c, \check{c}\}\}.$$

The partial map $\phi_3 : A_4{}^* \to A_3{}^*$ is defined by :

$$\mathrm{Dom}(\phi_3) = R_n$$

and for every $w = z_1 z_2 \dots z_h$, where $z_i \in A_4$,

$\phi_3(w) = \pi(z_{n+2})\pi(z_{2n+3}) \dots \pi(z_{k.(n+1)+1})$, with $k.(n+1)+1 < h \leq (k+1)(n+1)+1$.

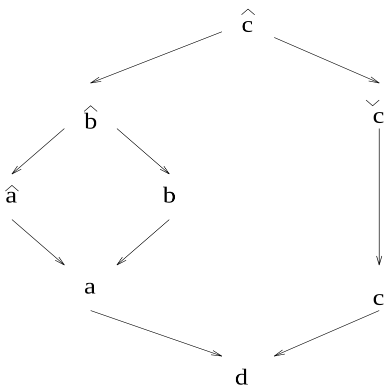

Figure 2: Lattice A_4

3.1.5 System S_5

Let $A_5 = \{x, \bar{x}, y\}$. Let us note $u_1 = x\bar{x}, u_2 = x^2\bar{x}^2$.

Let $\psi : A_4{}^* \to A_5{}^*$ be the unique monoid homomorphism such that:

$$\begin{array}{rrr} \psi(\hat{b}) = yu_2u_2u_2u_2, & \psi(b) = yu_1u_1u_2u_2, & \psi(\hat{a}) = yu_2u_2u_2, \\ \psi(a) = yu_1u_1u_2, & \psi(\hat{c}) = yu_2u_2u_2u_2u_1, & \psi(\check{c}) = yu_2u_2u_1u_1u_1, \\ \psi(c) = yu_1u_1u_1u_1u_1, & \psi(d) = y. & \end{array} \tag{13}$$

The map $\tau_4 : A_4{}^* \to A_5{}^*$ is then defined by:

$$\forall w \in A_4{}^*, \tau_4(w) = \psi(w)y.$$

The system S_5 consists of the rules

$$\begin{array}{rcl} \psi(L)y & \longrightarrow & \psi(M)y \\ u_2u_2u_2\psi(d^ncd^n)yu_1u_1 & \longrightarrow & u_2u_2u_2\psi(b^n)yu_2u_2 \\ x\bar{x} & \longrightarrow & \epsilon \end{array} \tag{14}$$

We denote by $\mathcal{D}$ the semi-Thue system consisting of the single rule :

$$x\bar{x} \longrightarrow \epsilon.$$

The set of words $[\epsilon]_{\overset{*}{\longleftrightarrow}_{\mathcal{D}}}$ is denoted by D_1^*, it is the so-called Dyck-language. Let us notice that the ordered set $(A_4, \overset{*}{\longrightarrow}_{\bar{S_4}})$ is a lattice (see its Hasse diagram on figure 2).

Let us denote by L_4 the set $\psi(A_4)$. One can check that $(L_4, \xrightarrow{*}_{\mathcal{D}})$ is a lattice too and that $\psi : A_4 \rightarrow L_4$ is a lattice-isomorphism. We define a map $\bar{\psi} : yD_1^* \rightarrow A_4^*$ by:

$$\forall w \in yD_1^*, \bar{\psi}(w) = \psi^{-1}(\bigwedge_{L_4}(\Delta_{\mathcal{D}}^*(w) \cap L_4)) \tag{15}$$

(where the symbol $\bigwedge_{L_4}$ denotes the g.l.b. in $(L_4, \xrightarrow{*}_{\mathcal{D}})$).
As yD_1^* is a suffix code which is the base of $(yD_1^*)^*$, $\bar{\psi}$ admits a unique extension as a morphism $(yD_1^*)^* \rightarrow A_4^*$ which is still denoted $\bar{\psi}$. The map $\phi_4 : A_5^* \rightarrow A_4^*$ is then defined by:

$$\mathrm{Dom}(\phi_4) = (yD_1^*)^* y$$

$$\forall w \in (yD_1^*)^* y, \phi_4(w) = \bar{\psi}(wy^{-1}).$$

3.2 General properties

3.2.1 Encodings

Proposition 3.1 *For every $w, w' \in A_0^*$,*
$w \longrightarrow_{S_0} w' \Longrightarrow \tau_0(w) \longrightarrow_{S_1} \tau_0(w')$.

Proof: Straightforward. □

Proposition 3.2 *For every $w, w' \in A_1^*$,*
$w \longrightarrow_{S_1} w' \Longrightarrow \tau_1(w) \longrightarrow_{S_2} \tau_1(w')$.

Proof: Straightforward. □

Proposition 3.3 *For every $w, w' \in A_2^*$,*
$w \longrightarrow_{S_2} w' \Longrightarrow \tau_3 \circ \tau_2(w) \xrightarrow{+}_{S_4} \tau_3 \circ \tau_2(w')$.

Proof: Let $w = \alpha u \beta, w' = \alpha v \beta$ and $(u, v) \in S_2$.
Then

$$\tau_3 \circ \tau_2(w) = \varphi(a\alpha)\varphi(u)\varphi(\beta)c, \quad \tau_3 \circ \tau_2(w') = \varphi(a\alpha)\varphi(v)\varphi(\beta)c.$$

case 1 : $u = \tau_1(u_i) = L_i, v = \tau_1(v_i)$ (where $0 \leq i \leq n$)
There exist $z \in \{a, b\}, L_i' \in \{a, b\}^*, t \in \{\hat{a}, \hat{b}, c\}, \gamma', \gamma'' \in A_4^*$ such that

$$\varphi(a\alpha) = \gamma' \hat{z} b^n, L_i = L_i' b, \varphi(\beta)c = t\gamma'' \tag{16}$$

Then

$$\tau_3 \circ \tau_2(w) = \gamma' \hat{z} b^n \varphi(L'_i) \hat{b} b^n t \gamma'' = \gamma' \hat{z} b^{n-i} (b^i \varphi(L'_i) \hat{b} b^{n-i}) b^i t \gamma'' \tag{17}$$

One can check that

$$b^i \varphi(L'_i) \hat{b} b^{n-i} \xrightarrow{*}_{S_4} L \tag{18}$$

because

- $|b^i \varphi(L'_i) \hat{b} b^{n-i}| = (l+1)(n+1) = |L|$,
- the image by π of the letters at positions $\equiv i+1 \pmod{n+1}$ in $b^i \varphi(L'_i) \hat{b} b^{n-i}$ (resp. in L) is L_i,
- every letter at position $j \not\equiv i+1 \pmod{n+1}$ in $b^i \varphi(L'_i) \hat{b} b^{n-i}$ (resp. in L) is equal to b (resp. belongs to $\{a, b\}$).

Hence

$$\begin{array}{lll} \gamma' \hat{z} b^{n-i} (b^i \varphi(L'_i) \hat{b} b^{n-i}) b^i t \gamma'' & \xrightarrow{*}_{S_4} & \gamma' \hat{z} b^{n-i} (L) b^i t \gamma'' \\ & & \text{(using rule } b \longrightarrow a \text{ as many times as necessary)} \\ & \longrightarrow_{S_4} & \gamma' \hat{z} b^{n-i} (M) b^i t \gamma''. \end{array} \tag{19}$$

Let $M'_i \in A_4^*$ be such that $M_i = M'_i c$. Let us define an homomorphism $\theta : {A_3}^* \to {A_4}^*$ (which is analogous to φ) by:

$$\theta(a) = ad^n, \;\; \theta(b) = bd^n, \;\; \theta(c) = cd^n, \;\; \theta(d) = dd^n.$$

One can notice that, by arguments similar to those used for derivation (18),

$$M \xrightarrow{*}_{S_4} d^i \theta(M'_i) c d^{n-i} \tag{20}$$

Hence

$$\begin{array}{lll} \gamma' \hat{z} b^{n-i} (M) b^i t \gamma'' & \xrightarrow{*}_{S_4} & \gamma' \hat{z} b^{n-i} (d^i \theta(M'_i) c d^{n-i}) b^i t \gamma'' \\ & \xrightarrow{*}_{S_4} & \gamma' \hat{z} d^n \theta(M'_i) c d^n t \gamma''. \end{array} \tag{21}$$

Owing to the equality $M_i = \gamma(\tau_1(v_i) c^{6(\nu - |v_i|)}) c$ and using rules 9-14 of (11), we get:

$$\hat{z} d^n \theta(M'_i) c d^n t \xrightarrow{*}_{S_4} \varphi(z \tau_1(v_i)) t \tag{22}$$

By all the above derivations we have:

$$\begin{array}{lll} \tau_3 \circ \tau_2(w) & \longrightarrow_{S_4} & \gamma' \hat{z} b^n \varphi(\tau_1(v_i)) t \gamma'' \\ & = & \varphi(a\alpha) \varphi(\tau_1(v_i)) \varphi(\beta) c \\ & = & \tau_3 \circ \tau_2(w'). \end{array} \tag{23}$$

case 2 : $u = b, v = a$.
Then $\tau_3 \circ \tau_2(w) \longrightarrow_{S_4} \tau_3 \circ \tau_2(w')$ (using rule $\hat{b} \longrightarrow \hat{a}$). $\square$

Proposition 3.4 *For every $w, w' \in A_4^*$,*
$w \longrightarrow_{S_4} w' \Longrightarrow \tau_4(w) \xrightarrow{+}_{S_5} \tau_4(w')$.

Proof: Not difficult. □

3.2.2 Decodings

Proposition 3.5 *Let $w, w' \in A_5^*$ such that $w \in (yD_1^*)^* y$ and*

$$w \longrightarrow_{S_5} w'. \tag{24}$$

Then

1. $w' \in (yD_1^*)^* y$
2. $\phi_4(w) \xrightarrow{*}_{\bar{S}_4} \phi_4(w')$
3. *If the rule used in derivation (24) is $\psi(L)y \longrightarrow \psi(M)y$ then* $\phi_4(w) \longrightarrow_{S_4} \phi_4(w')$

Proof: Let us notice first that, for every $w \in L_4$, the equality (15) implies that $\bar{\psi}(w) = \psi^{-1}(w)$. Hence

$$\bar{\psi} \circ \psi = \mathrm{Id}_{A_4^*}. \tag{25}$$

Let $w, w' \in A_5^*$ such that $w \in (yD_1^*)^* y$ and $w \longrightarrow_{S_5} w'$.
case 1 : The rule used is $\psi(L)y \longrightarrow \psi(M)y$
The occurrence of $\psi(L)y$ in w must be of the form

$$w = \alpha\psi(L)y\beta \text{ with } \alpha \in (yD_1^*)^* \text{ and } \ \beta \in (D_1^* y)^*.$$

Hence

$$w' = \alpha\psi(M)y\beta$$

which shows that $w' \in (yD_1^*)^* y$ (point (1) of the proposition), and

$$\phi_4(w) = \bar{\psi}(\alpha)\bar{\psi}(\psi(L))\phi_4(y\beta), \ \ \phi_4(w') = \bar{\psi}(\alpha)\bar{\psi}(\psi(M))\phi_4(y\beta).$$

By identity (25)

$$\phi_4(w) = \bar{\psi}(\alpha)L\phi_4(y\beta), \ \ \phi_4(w') = \bar{\psi}(\alpha)M\phi_4(y\beta).$$

Hence $\phi_4(w) \longrightarrow_{S_4} \phi_4(w')$ (proving points (2)(3) of the proposition).
case 2 : The rule used is $u_2u_2u_2\psi(d^ncd^n)yu_1u_1 \longrightarrow u_2u_2u_2\psi(b^n)yu_2u_2$
Then

$$w = \alpha y w_1 u_2u_2u_2\psi(d^ncd^n)yu_1u_1w_2\beta y$$

$$w' = \alpha y w_1 u_2 u_2 u_2 \psi(b^n) y u_2 u_2 w_2 \beta y$$

with $\alpha, \beta \in (yD_1^*)^*$, $w_1, w_2 \in D_1^*$.
As $\psi(\hat{a}) \in \Delta^*_{\mathcal{D}}(yw_1u_2u_2u_2)$

$$\bar{\psi}(yw_1u_2u_2u_2) \xrightarrow{*}_{\mathcal{D}} \hat{a}.$$

Hence $\bar{\psi}(yw_1u_2u_2u_2) \in \{\hat{a}, \hat{b}, \hat{c}\}$.
Let us examine all the possible values of the pair $(\bar{\psi}(yu_1u_1w_2), \bar{\psi}(yu_2u_2w_2))$.
One can check that for every $v \in yD_1^*$, the following cases are possible:

$$\begin{array}{lll}
\Delta^*_{\mathcal{D}}(v) \cap \{\psi(a), \psi(b), \psi(c)\} = \emptyset, & \bar{\psi}(v) = d & \textbf{(C1)} \\
\Delta^*_{\mathcal{D}}(v) \cap \{\psi(a), \psi(b), \psi(c)\} = \{\psi(a)\}, & \bar{\psi}(v) \in \{a, \hat{a}\} & \textbf{(C2)} \\
\Delta^*_{\mathcal{D}}(v) \cap \{\psi(a), \psi(b), \psi(c)\} = \{\psi(a), \psi(b)\}, & \bar{\psi}(v) \in \{b, \hat{b}\} & \textbf{(C3)} \\
\Delta^*_{\mathcal{D}}(v) \cap \{\psi(a), \psi(b), \psi(c)\} \supseteq \{\psi(c)\}, & \bar{\psi}(v) \in \{c, \check{c}, \hat{c}\} & \textbf{(C4)}.
\end{array}$$

But the fact that every word in $\{\psi(a), \psi(b), \psi(c)\}$ begins with yu_1u_1 implies that

$$\Delta^*_{\mathcal{D}}(yu_1u_1w_2) \cap \{\psi(a), \psi(b), \psi(c)\} = \Delta^*_{\mathcal{D}}(yu_2u_2w_2) \cap \{\psi(a), \psi(b), \psi(c)\}.$$

Hence both words of $yu_1u_1w_2, yu_2u_2w_2$ fufill the same case Ci $(1 \leq i \leq 4)$.
It follows that

$$(\bar{\psi}(yu_1u_1w_2), \bar{\psi}(yu_2u_2w_2)) \in \{(x,x) | x \in A_4\} \cup \{(a, \hat{a}), (b, \hat{b}), (c, \check{c}), (c, \hat{c}), (\check{c}, \hat{c})\}.$$

Let us prove that the two last values $(c, \hat{c}), (\check{c}, \hat{c})$ are impossible.

$$\begin{array}{rll}
\bar{\psi}(yu_2u_2w_2)) = \hat{c} \Longleftrightarrow & \{\psi(a), \psi(c)\} \subseteq \Delta^*_{\mathcal{D}}(yu_2u_2w_2) & \textbf{(C5)} \\
& \text{or } \{\psi(\hat{c})\} \subseteq \Delta^*_{\mathcal{D}}(yu_2u_2w_2) & \textbf{(C6)}.
\end{array}$$

- in case (C5), $\{\psi(a), \psi(c)\} \subseteq \Delta^*_{\mathcal{D}}(yu_1u_1w_2)$ too, so that

$$\bar{\psi}(yu_1u_1w_2) = \hat{c}$$

- in case (C6), $yu_1u_1u_2u_2u_1 \in \Delta^*_{\mathcal{D}}(yu_1u_1w_2)$, hence $\{\psi(a), \psi(c)\} \subseteq \Delta^*_{\mathcal{D}}(yu_1u_1w_2)$ and again

$$\bar{\psi}(yu_1u_1w_2) = \hat{c}$$

It follows that

$$(\bar{\psi}(yu_1u_1w_2), \bar{\psi}(yu_2u_2w_2)) \in \{(x,x) | x \in A_4\} \cup \{(a, \hat{a}), (b, \hat{b}), (c, \check{c})\}.$$

Hence $\phi_4(w) \xrightarrow{*}_{\bar{S}_4} \phi_4(w')$ using no rule or one of the rules 9-14 of (11) or 3-5 of (12).
case 3 : The rule used is $x\bar{x} \longrightarrow \epsilon$.

Let $w_1, w'_1 \in A_5^*$ such that $w = w_1 y, w' = w'_1 y$.
$w_1 \xrightarrow{*}_{\mathcal{D}} w'_1$ hence $\Delta^*_{\mathcal{D}}(w_1) \supseteq \Delta^*_{\mathcal{D}}(w'_1)$, so that

$$\bigwedge_{L_4}(\Delta^*_{\mathcal{D}}(w_1) \cap L_4) \xrightarrow{*}_{\mathcal{D}} \bigwedge_{L_4}(\Delta^*_{\mathcal{D}}(w'_1) \cap L_4)$$

As ψ is a lattice isomorphism fromA_4 to L_4 we have

$$\psi^{-1}(\bigwedge_{L_4}(\Delta^*_{\mathcal{D}}(w_1) \cap L_4)) \xrightarrow{*}_{\bar{S_4}} \psi^{-1}(\bigwedge_{L_4}(\Delta^*_{\mathcal{D}}(w'_1) \cap L_4)).$$

□

Proposition 3.6 *Let $w, w' \in R_n$ such that $w \longrightarrow_{\bar{S_4} - \{\hat{c} \longrightarrow \hat{b}\}} w'$.*
Then $\phi_3(w) \xrightarrow{*}_{\bar{S_3}} \phi_3(w')$.

Proof: Easy. □

Proposition 3.7 *Let $w, w' \in A_3^*$ such that $w \longrightarrow_{\bar{S_3}} w'$.*
Then $\phi_2(w) \xrightarrow{*}_{S_2} \phi_2(w')$.

Proof: Easy. □

Proposition 3.8 *Let $w, w' \in A_2^*$ and $\bar{w} \in \mathrm{Im}\tau_1$ such that*
$\bar{w} \xrightarrow{*}_{\mathcal{B}} w \longrightarrow_{S_2} w'$.
Then $\exists i \in \{0, 1\}, \exists \bar{w}' \in \mathrm{Im}\tau_1$ such that $\bar{w} \xrightarrow{i}_{S_2 - \mathcal{B}} \bar{w}' \xrightarrow{*}_{\mathcal{B}} w'$.

Proof: Easy. □

Proposition 3.9 *Let $w, w' \in A_2^*$ such that $w \longrightarrow_{S_2 - \mathcal{B}} w'$.*
Then $\phi_1(w) \longrightarrow_{S_1} \phi_1(w')$.

Proof: Not difficult. □

Proposition 3.10 *Let $w, w' \in A_1^*$ such that $w \longrightarrow_{S_1} w'$. Then*

1. *if the rule used is $\eta(e^l) \longrightarrow \epsilon$ then $\phi_0(w) = \phi_0(w')$*
2. *if the rule used is not $\eta(e^l) \longrightarrow \epsilon$ then*
   $\phi_0(w) \longrightarrow_{S_0} \phi_0(w')$.

Proof: Not difficult. □

3.2.3 Stability

Proposition 3.11 *Let $w, w' \in A_5^*$ such that $w \in \Delta_{\mathcal{D}}^*(\mathrm{Im}\tau_4)$ and $w \longrightarrow_{S_5} w'$. Then $w' \in \Delta_{\mathcal{D}}^*(\mathrm{Im}\tau_4)$.*

Proof: Straightforward. □

Proposition 3.12 *Let $w, w' \in A_4^*$ such that $w \in R_n$ and $w \longrightarrow_{\bar{S}_4} w'$. Then $w' \in R_n$.*

Proof: Easy. □

3.2.4 Extractions

Proposition 3.13 *Let $w \in A_5^*$ such that $w \stackrel{\infty}{\longrightarrow}_{S_5}$. Then, $\exists w' \in (yD_1^*)^* y, w' \stackrel{\infty}{\longrightarrow}_{S_5}$.*

Sketch of proof: Let us define a morphism $\phi_5' : (y(x+\bar{x}+y)^*)^* \longrightarrow (yD_1^*)^*$ as follows:
let w be some word in $y(x+\bar{x}+y)^*$. It has a unique decomposition as

$w = yw_0z_1w_1 \dots z_iw_i \dots z_kw_k$ where $w_i \in D_1^*, z_i \in \{x, \bar{x}\}$ and $z_1 \dots z_i \dots z_k = \rho_{\mathcal{D}}(w)$.

Then

$$\phi_5'(w) = yw_0w_1 \dots w_i \dots w_k.$$

We then define a map $\phi_5 : A_5^* \longrightarrow (yD_1^*)^* y$ by

$$\forall w \in A_5^*, \;\; \phi_5(w) = \phi_5'(ywy).$$

One can check that for all $w, w' \in A_5^*$,

$$w \longrightarrow_{S_5} w' \Longrightarrow \phi_5(w) \longrightarrow_{S_5} \phi_5(w').$$

It follows that

$$w \stackrel{\infty}{\longrightarrow}_{S_5} \Longrightarrow \phi_5(w) \stackrel{\infty}{\longrightarrow}_{S_5} .$$

□

Proposition 3.14 *Let $w \in A_4^*$ such that $w \stackrel{\infty}{\longrightarrow}_{\bar{S}_4}$. Then, $\exists w' \in R_n, w' \stackrel{\infty}{\longrightarrow}_{\bar{S}_4}$.*

Sketch of proof: Let $D = (w_i)_{i \in \mathbb{N}}$ be an infinite derivation $\pmod{\bar{S}_4}$ starting from $w = w_0$.
Let us call *main* letters the elements of $\{\hat{a}, \hat{b}, \hat{c}\}$ and *nominal* letters the elements of $\{a, b, c, \check{c}\}$. For every $v \in A_4^*$ we call *alternation* any factor f of v of the form:

$f = z_1 g z_2$ such that z_1, z_2 are main letters , $g \in A_4^*$ contains only nominal letters,

$$\text{and} \quad | g | \not\equiv n \pmod{n+1}.$$

We denote by $\|v\|$ the *number of alternations* of v. We observe first that relation $\longrightarrow_{\bar{S}_4}$ decreases the number of alternations. Without loss of generality we can the suppose that all w_i have the same number of alternations, say p. By left-product by a fixed main letter, we can also suppose that all w_i are beginning with the main letter $\hat{a}$. Let us consider the decompositions

$$w_i = z_{i,0} w_{i,0} z_{i,1} w_{i,1} \dots z_{i,j} w_{i,j} \dots z_{i,p} w_{i,p}$$

where

$z_{i,j}$ is nominal , $z_{i,j} w_{i,j}$ has no alternation and, $z_{i,j} w_{i,j} z_{i,j+1}$ has one alternation .

$z_{i,j} w_{i,j}$ will be called the $(j+1)$-th *block* of w_i. By lemma 2.2 , we can suppose that D is a rl-derivation. We show now that either $p = 0$ or there exists another infinite derivation with at most $p-1$ alternations. Let us suppose $p > 0$.

- If some step of derivation applies one of the rules $\hat{b} \longrightarrow b, \hat{a} \longrightarrow a, \hat{c} \longrightarrow \check{c}$ on $z_{i,p}$, then no rule applied later can involve any position of the $p+1$-th block. Hence, $(w_{i+1+k}(z_{i+1,p} w_{i+1,p})^{-1})_{k \in \mathbb{N}}$ is an infinite derivation with $p-1$ alternations.
- If some step of derivation applies one rule in some j-th block ($1 \leq j \leq p$), then no rule applied later can involve any position of the $p+1$-th block. Hence we obtain again some infinite derivation with p-1 alternations.
- if none of the two above cases is fulfilled, then every rule applies on the $p+1$-th block, and the prefix $z_{i,0} w_{i,0} z_{i,1} w_{i,1} \dots z_{i,p-1} w_{i,p-1}$ is fixed. Hence $(z_{i,p} w_{i,p})_{i \in \mathbb{N}}$ is an infinite derivation with 0 alternation.

We have proved by induction that, under the hypothesis of the proposition, there exists some infinite derivation $D' = (w'_i)_{i \in \mathbb{N}}$ with 0 alternations and, in fact, in Q_n. Let $r \in [1, n+1]$ such that $|w'_i| \equiv r \pmod{n+1}$ (for all $i \in \mathbb{N}$). Then $(\hat{a} b^n w'_i b^{n+1-r} \check{c})_{i \in \mathbb{N}}$ is an infinite derivation starting in R_n. $\square$

Proposition 3.15 *Let $w \in A_2^*$ such that $w \overset{\infty}{\longrightarrow}_{S_2}$. Then,*

1. $\exists w' \in A_2^*, w' \overset{\infty}{\longrightarrow}_{S_2 - \mathcal{B}}$

2. *if* $w \in \mathrm{Im}\tau_1, w \xrightarrow{\infty}_{S_2-\mathcal{B}}$.

Sketch of proof: Let us suppose that $w \in A_2^*$ is such that $w \xrightarrow{\infty}_{S_2}$. Let us prove point (1).
We observe that the number of factors b^3 can only decrease in a derivation (mod S_2) . Hence there exists $w_1 \in A_2^*$ such that $w_1 \xrightarrow{\infty}_{S_2}$ and the number of factors b^3 is fixed (throughout the derivation).
Now, in such a derivation, the number of consecutive blocks b^2 at distance < 4 (i.e. factors of the form b^2ub^2 with no occurence of b^2 in bub and $|u| < 4$) can only decrease. Hence there exists $w_2 \in A_2^*$ such that $w_2 \xrightarrow{\infty}_{S_2}$ and the number of factors b^3 is fixed and the number of consecutive blocks b^2 at distance < 4 is fixed (throughout the derivation).
In such a derivation the rule $b \longrightarrow a$ can be applied only:

- between two consecutive b^2 at distance ≥ 5
- or before the first block b^2 or after the last block b^2
- or inside a block b^2
- or can transform a block b^2abaab^2 or b^2aabab^2 into b^2aaaab^2

In either case the position of the new letter a introduced by the rule will never be usable later on in the derivation. Hence we can discard all the applications of the rule $b \longrightarrow a$ in the derivation, and obtain another infinite derivation (mod $S_2 - \mathcal{B}$) starting on w_2. Point (1) is proved.
If $w \in \mathrm{Im}\tau_1$, then w has no factor b^3 nor consecutive b^2 at distance < 4. Hence, by the above arguments, $w \xrightarrow{\infty}_{S_2-\mathcal{B}}$. Point (2) is proved. □

4 Reductions and bounds

Let us use the notation:

$$\tau = \tau_4 \circ \tau_3 \circ \tau_2 \circ \tau_1 \circ \tau_0 \quad , \quad \phi = \phi_0 \circ \phi_1 \circ \phi_2 \circ \phi_3 \circ \phi_4.$$

An easy verification from the definitions shows that

$$\forall 0 \leq i \leq 4, \quad \phi_i \circ \tau_i = \mathrm{Id}_{A_i^*}. \tag{26}$$

which also implies that

$$\phi \circ \tau = \mathrm{Id}_{A_0^*}. \tag{27}$$

4.1 Infinitary problems

4.1.1 Termination problem

Theorem 4.1 *There exists some semi-Thue system S with 3 rules and with undecidable termination problem.*

Proof: Let S_0 be some finite semi-Thue system over some finite alphabet A_0 and with undecidable termination problem (for example we can take $S_0 = \mathcal{T}$, where $\mathcal{T}$ is the system defined in §2.2.2). Let us consider the sequence of systems $(S_i)_{0 \leq i \leq 5}$ defined in section 3.1 and starting by the above S_0. We reduce now the termination problem for S_0 to the termination problem for S_5 by showing that

$$\forall u \in A_0^*, u \xrightarrow{\infty}_{S_0} \Longleftrightarrow \tau(u) \xrightarrow{\infty}_{S_5} . \tag{28}$$

It follows that the system $S = S_5$ has the required property.
Let $u \in A_0^*$.

1- Let us suppose $u \xrightarrow{\infty}_{S_0}$.
By Propositions 3.1, 3.2, 3.3, 3.4, we obtain

$$\tau(u) \xrightarrow{\infty}_{S_5} .$$

2- Let us suppose

$$\tau(u) \xrightarrow{\infty}_{S_5} . \tag{29}$$

By Proposition 3.5 and owing to the fact that $S_5 - \{\psi(L)y \longrightarrow \psi(M)y\}$ is u-terminating, we have

$$\phi_4(\tau(u)) \xrightarrow{\infty}_{\bar{S}_4} .$$

By Proposition 3.11 the whole derivation (29) lies inside $\Delta^*_{\mathcal{D}}(\mathrm{Im}\tau_4)$, hence

$$\phi_4(\tau(u)) \xrightarrow{\infty}_{\bar{S}_4 - \{\hat{c} \longrightarrow \hat{b}\}} .$$

By Proposition 3.6 and owing to the fact that $\bar{S}_4 - \{L \longrightarrow M\}$ is u-terminating,

$$\phi_3(\phi_4(\tau(u))) \xrightarrow{\infty}_{\bar{S}_3} .$$

By Proposition 3.7 and owing to the fact that $\{b \longrightarrow a, c \longrightarrow \epsilon, a \longrightarrow d\}$ is u-terminating,

$$\phi_2(\phi_3(\phi_4(\tau(u)))) \xrightarrow{\infty}_{S_2} .$$

Using identities (26) we observe that

$$\phi_2(\phi_3(\phi_4(\tau(u)))) = \tau_1(u) \in \mathrm{Im}\tau_1.$$

Hence by Proposition 3.15, point (2),

$$\phi_2(\phi_3(\phi_4(\tau(u)))) \xrightarrow{\infty}_{S_2 - \mathcal{B}}$$

and by Propositions 3.9, 3.10

$$\phi_0(\phi_1(\phi_2(\phi_3(\phi_4(\tau(u)))))) \xrightarrow{\infty}_{S_0}$$

i.e, by identity (27)

$$u \xrightarrow{\infty}_{S_0} .$$

□

4.1.2 Uniform termination problem

Theorem 4.2 *The uniform termination problem is undecidable for 3 rules semi-Thue systems.*

Proof: We reduce below the uniform termination problem for finite semi-Thue systems to the uniform termination problem for 3 rules semi-Thue systems. As the former is undecidable (see §2.2.1), the second problem is undecidable too.
Let S_0 be some finite semi-Thue system over some finite alphabet A_0. Here again we consider the sequence of systems $(S_i)_{0 \leq i \leq 5}$ defined in section 3.1 and starting by the above S_0. We show now that the uniform termination property for S_0 is equivalent to the uniform termination property for S_5:

$$\exists u \in A_0^*, u \xrightarrow{\infty}_{S_0} \Longleftrightarrow \exists w \in A_5^*, w \xrightarrow{\infty}_{S_5} . \tag{30}$$

1- Let us suppose $u \xrightarrow{\infty}_{S_0}$.
By the same arguments as in part 1 of the proof of Proposition 4.1, we then have

$$\tau(u) \xrightarrow{\infty}_{S_5} .$$

2- Let us suppose that there exists some $w \in A_5^*$ such that

$$w \xrightarrow{\infty}_{S_5} .$$

By Proposition 3.13

$$\exists w' \in (yD_1^*)^* y, w' \xrightarrow{\infty}_{S_5} .$$

By Proposition 3.5

$$\phi_4(w') \xrightarrow{\infty}_{\bar{S_4}} .$$

By Proposition 3.14

$$\exists w'' \in R_n, \;\; w'' \xrightarrow{\infty}_{\bar{S_4}} .$$

Owing to the fact that $\longrightarrow_{\bar{S_4}}$ decreases the number of $\hat{c}$,

$$\exists w''' \in R_n, \;\; w'' \xrightarrow{*}_{\bar{S_4}} w''' \xrightarrow{\infty}_{\bar{S_4} - \{\hat{c} \longrightarrow \hat{b}\}}$$

and by Proposition 3.12, the whole derivation $w''' \longrightarrow^{\infty}_{\bar{S_4}-\{\hat{c}\longrightarrow\hat{b}\}}$ is inside R_n. Using now Proposition 3.6 and the fact that $S_4 - \{L \longrightarrow M\}$ is u-terminating, we obtain

$$\phi_3(w''') \longrightarrow^{\infty}_{\bar{S_3}} .$$

Using Proposition 3.7 and the fact that $\{a \longrightarrow d, c \longrightarrow \epsilon, b \longrightarrow a\}$ is u-terminating, we obtain

$$\phi_2(\phi_3(w''')) \longrightarrow^{\infty}_{S_2} .$$

By Proposition 3.15, point (1)

$$\exists w'''' \in A_2^*, \quad w'''' \longrightarrow^{\infty}_{S_2-\mathcal{B}}$$

which, by Propositions 3.9,3.10 implies

$$\phi_0(\phi_1(w'''')) \longrightarrow^{\infty}_{S_0} .$$

We have then exhibited some $u \in A_0^*$ such that $u \longrightarrow^{\infty}_{S_0}$, as was required. Equivalence (30) is thus proved. □

4.2 Finitary problems

4.2.1 Accessibility and Common descendant problems

Theorem 4.3 *There exist some semi-Thue system S with 3 rules which has undecidable individual accessibility problem (for some word w_0) and also undecidable common descendant problem.*

Proof: Let S_0 be some finite semi-Thue system over some finite alphabet A_0 such that it is undecidable, for words $u \in A_0^*$, whether $u \longrightarrow^{*}_{S_0} \epsilon$ or not. For example, we can take $A_0 = \{a, b, c, d, e, f\}$ and $S_0 = \mathcal{C} \cup \{faaaf \longrightarrow \epsilon\}$ (where $\mathcal{C}$ is the system given in §2.2.3).

Let us consider the sequence of systems $(S_i)_{0\leq i\leq 5}$ defined in section 3.1 and starting by the above S_0. We shall show that, for every $u \in A_0^*$, the 3 following statements are equivalent:

(1) $u \longrightarrow^{*}_{S_0} \epsilon$

(2) $\tau(u) \longrightarrow^{*}_{S_5} y^{3(n+1)+2}$

(3) $\tau(u), y^{3(n+1)+2}$ have some common descendant $\pmod{S_5}$.

(1) $\Longrightarrow$ (2):
Let us suppose $u \longrightarrow^{*}_{S_0} \epsilon$. Using Proposition 3.1 and the first rule of S_1 we obtain

$$\tau_0(u) \longrightarrow^{*}_{S_1} \eta(e^{l_0}) \longrightarrow_{S_1} \epsilon.$$

By Propositions 3.2, 3.3, 3.4

$$\tau(u) \xrightarrow{*}_{S_5} \tau_4(\tau_3(\tau_2(\tau_1(\epsilon)))) \tag{31}$$

where

$$\tau_4(\tau_3(\tau_2(\tau_1(\epsilon)))) = \tau_4(\hat{a}b^n\hat{b}b^n\hat{b}b^nc) = \psi(\hat{a}b^n\hat{b}b^n\hat{b}b^nc)y.$$

As for every $z \in A_4, \psi(z) \xrightarrow{*}_{\mathcal{D}} y$ we get

$$\tau_4(\tau_3(\tau_2(\tau_1(\epsilon)))) \xrightarrow{*}_{S_5} y^{3(n+1)+2}. \tag{32}$$

By derivations (31, 32), $\tau(u) \xrightarrow{*}_{S_5} y^{3(n+1)+2}$.
(2) $\Longrightarrow$ (3): this is obvious.
(3) $\Longrightarrow$ (1):
Let us suppose that $\tau(u), y^{3(n+1)+2}$ have some common descendant $\pmod{S_5}$.
As $y^{3(n+1)+2}$ is irreducible $\pmod{S_5}$, we can conclude that

$$\tau(u) \xrightarrow{*}_{S_5} y^{3(n+1)+2}.$$

By Proposition 3.5 we obtain

$$\phi_4(\tau(u)) \xrightarrow{*}_{\bar{S_4}} d^{3(n+1)+1}.$$

As $\phi_4(\tau(u)) \in \mathrm{Im}\tau_3$, $\phi_4(\tau(u)) \in R_n$ and has no occurence of $\hat{c}$. Hence by Proposition 3.6

$$\phi_3(\phi_4(\tau(u))) \xrightarrow{*}_{\bar{S_3}} dd$$

which, by Proposition 3.7 leads to

$$\phi_2(\phi_3(\phi_4(\tau(u)))) \xrightarrow{*}_{S_2} aa.$$

Applying Proposition 3.8 inductively, we obtain

$$\exists \bar{w}' \in \mathrm{Im}\tau_1, \phi_2(\phi_3(\phi_4(\tau(u)))) \xrightarrow{*}_{S_2-\mathcal{B}} \bar{w}' \xrightarrow{*}_{\mathcal{B}} aa.$$

The only possible value of $\bar{w}'$ is bb, hence

$$\phi_2(\phi_3(\phi_4(\tau(u)))) \xrightarrow{*}_{S_2-\mathcal{B}} bb$$

and Propositions 3.9, 3.10 then show that

$$\phi(\tau(u)) \xrightarrow{*}_{S_0} \epsilon \text{ i.e. } u \xrightarrow{*}_{S_0} \epsilon.$$

The equivalence between points (1),(2) and (3) is then established.
The fact that (1) $\Longleftrightarrow$ (2) proves that S_5 has undecidable individual accessibility problem for $w_0 = y^{3(n+1)+2}$.
The fact that (1) $\Longleftrightarrow$ (3) proves that S_5 has undecidable common descendant problem .
□

Corollary 4.4 *The Post correspondence problem is undecidable for 7 pairs of words.*

Proof: Follows from theorems 4.3 and 2.1. □
□

References

[1] S.I. Adjan and G.S. Makanin. Investigations on algorithmic problems in algebra. *Proceedings of the Steklov Institute of Mathematics, issue 3*, pages 207–226, 1986.

[2] M. Billaud, P. Lafon, Y. Métivier, and E. Sopena. Graph rewriting systems with priorities. In *Proceedings WG 89*, pages 94–106. LNCS 411, Springer-Verlag, 1989.

[3] R.V. Book and F. Otto. *String Rewriting Systems. Texts and monographs in Computer Science*. Springer-Verlag, 1993.

[4] W.W. Boone, D. Collins, and Y.V. Matiyasevich. Embedding into semigroups with only a few defining relations. In *Proceedings of the 2nd Scandinavian Logic Symposium*, pages 27–40. Studies in Logic and the foundations of Mathematics, 1971.

[5] V.V. Borisov. Simple examples of groups with unsolvable word problem. *Mat.Zametki 6, english translation in Math. notes 6*, pages 521–532, 1969.

[6] A.C. Caron. Linear bounded automata and rewrite systems: influence of initial configurations on decision properties. In *Proceedings CAAP 91*, pages 74–89. LNCS Springer-Verlag, 1991.

[7] V. Claus. Some remarks on PCP(k) and related problems. *Bull. EATCS 12*, pages 54–61, 1980.

[8] D.J. Collins. A simple presentation of a group with unsolvable word problem. *Illinois J. Math., 30 (2)*, pages 230–234, 1986.

[9] M. Dauchet. Simulation of Turing machines by a regular rewrite rule. *TCS vol. 103*, pages 409–420, 1992.

[10] M. Davis. *Computability and Unsolvability*. Mc Graw-Hill, 1958.

[11] N. Dershowitz. Termination of rewriting. *Journal of Symbolic Computation no 3*, pages 69–116, 1987.

[12] N. Dershowitz and J.P. Jouannaud. Rewrite Systems. In *Handbook of theoretical computer science, vol.B, Chapter 2*, pages 243–320. Elsevier, 1991.

[13] A. Ehrenfeucht and G. Rozenberg. On the (generalized) Post correspondence problem with lists of length 2. *LNCS 115*, pages 408–416, 1981.

[14] P.K. Hooper. The undecidability of the Turing machine immortality problem. *The Journal of Symbolic Logic, vol.31,nr 2*, pages 219–234, 1966.

[15] G. Huet. Confluent reductions: abstract properties and applications to term rewriting systems. *JACM vol. 27 no 4*, pages 797–821, 1980.

[16] G. Huet and D. Lankford. *On the uniform halting problem for term rewriting systems.* Rapport Laboria, 1978.

[17] M. Jantzen. *Confluent String Rewriting (EATCS monograph).* Springer-Verlag, 1988.

[18] W. Kurth. Termination und konfluenz von semi-Thue-systemen mit nur einer regel. *Dissertation, Technische Universität Clausthal*, 1990.

[19] G. Lallement. The Word Problem for Thue Rewriting Systems. In *Term Rewriting,French Spring School of Theoretical Computer Science, Font-Romeu 93, Advanced course*, pages 27–38. Springer-Verlag, LNCS 909, 1994.

[20] M. Margenstern. Turing machines : on the frontier between a decidable halting problem and universality. In *Proceedings FCT'93*, pages 375–385. Springer-Verlag, LNCS 710, 1993.

[21] J.V. Matiyasevich. Simple examples of undecidable associative calculi. *Soviet Math Dokladi*, pages 555–557, 1967.

[22] J.V. Matiyasevich. Simple examples of unsolvable canonical calculi. *Trudy Mat. Inst. Steklov 93*, pages 61–110, 1967.

[23] J.V. Matiyasevich. On investigations on some algorithmic problems in algebra and number theory. *Proceedings of the Steklov Institute of Mathematics*, pages 227–252, 1986.

[24] Y. Matiyasevich. Word problem for Thue systems with a few relations. In *Term Rewriting,French Spring School of Theoretical Computer Science, Font-Romeu 93, Advanced course*, pages 39–53. Springer-Verlag,LNCS 909, 1994.

[25] Y. Matiyasevich and G. Sénizergues. Decision problems for semi-Thue systems with a few rules. In *Proceedings LICS96*, pages 523–531. IEEE Computer Society Press, 1996.

[26] R. McNaughton. The uniform halting problem for one-rule semi-Thue systems. *Report 94-18 of Rensselaer Polytechnic Institute*, 1994.

[27] A. Middeldorp and B. Gramlich. Simple termination is difficult. In *Proceedings RTA93*, pages 228–242. Springer-Verlag,LNCS 690, 1993.

[28] J.J. Pansiot. A note on Post's correspondence problem. *IPL vol. 12*, pages 233–233, 1981.

[29] G. Sénizergues. Some undecidable termination problems for semi-Thue systems. *TCS vol. 142*, pages 257–276, 1995.

[30] G. Sénizergues. On the termination-problem for one-rule semi-Thue systems. In *Proceedings RTA96, LNCS 1103*, pages 302–316. Springer-Verlag, 1996.

[31] K. Shikishima-Tsuji, M. Katsura, and Y. Kobayashi. On termination of confluent one0rule string-rewriting systems. *Information Processing Letters*, pages 91–96, 1997.

[32] H. Zantema and A. Geser. A complete characterisation of termination of $0^p1^q \rightarrow 1^r0^s$. In *Proceedings RTA95, LNCS 914*, pages 41–55. Springer-Verlag, 1995.