On One-Rule Grid Semi-Thue Systems
Michel Latteux, Yves Roos

To cite this version:
Michel Latteux, Yves Roos. On One-Rule Grid Semi-Thue Systems. Fundamenta Informaticae, Polskie Towarzystwo Matematyczne, 2012, Words, Graphs, Automata and Languages. Special Issue Honoring the 60th Birthday of Professor Tero Harju, 116 (1-4), pp.189-204. <10.3233/FI-2012-678>. <hal-00749289>

HAL Id: hal-00749289
https://hal.inria.fr/hal-00749289
Submitted on 7 Nov 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
On One-Rule Grid Semi-Thue Systems

Michel Latteux
Laboratoire d’Informatique Fondamentale de Lille
Université Lille 1
michel.latteux@lifl.fr

Yves Roos
Laboratoire d’Informatique Fondamentale de Lille
Université Lille 1
yves.roos@lifl.fr

Abstract. The family of one-rule grid semi-Thue systems, introduced by Alfons Geser, is the family of one-rule semi-Thue systems such that there exists a letter $c$ that occurs as often in the left-hand side as the right-hand side of the rewriting rule. We prove that for any one-rule grid semi-Thue system $S$, the set $S(w)$ of all words obtainable from $w$ using repeatedly the rewriting rule of $S$ is a constructible context-free language. We also prove the regularity of the set $\text{Loop}(S)$ of all words that start a loop in a one-rule grid semi-Thue systems $S$.

Keywords: one-rule semi-Thue system, termination, grid semi-Thue system.

1. Introduction

Semi-Thue systems, that are the non-symmetrical version of Thue systems introduced by Axel Thue in 1914, serve as a model for rewriting systems. Thus they are of primordial interest for computational problems. For several years they have been intensively studied and several deep results have been obtained [10, 3, 2, 21, 19, 22]. However some intriguing decidability problems remain open. With a semi-Thue system $S$, one associates the set $S_\infty$ of words that start an infinite derivation in $S$. The recursiveness of $S_\infty$ is called the termination problem for $S$ and the emptiness of $S_\infty$ is the uniform termination problem for $S$. The best result on the set $S_\infty$ is that the termination problem and the uniform termination problem are undecidable for 3-rules semi-Thue systems ([14]). Clearly the termination problem is decidable for length-preserving semi-Thue systems, contrarily to the uniform termination problem that has been shown undecidable for length-preserving semi-Thue systems.
In the original definition of Alfons Geser, a one-rule grid semi-Thue system is a one rule semi-Thue system 1

A M. Latteux, Y. Roos / On one-rule grid semi-Thue systems

having a letter terminating.

We show that the words S system: for instance a simple rational expression. This property can not be generalized to arbitrary 1-rule semi-Thue systems in 1

First, we give a new formulation of this decidable characterization of non-uniformly terminating length 1 or 2 was previously shown decidable by Winfried Kurth in [11].

to know whether a one-rule semi-Thue system has a loop of that are uniformly terminating by proving that such a system is non-uniformly terminating if and only if it has a loop of length 1 or 2. To know whether a one-rule semi-Thue system has a loop of length 1 or 2 was previously shown decidable by Winfried Kurth in [11].

One-rule semi-Thue systems are the simplest rewriting systems. Indeed, they are defined by two

words l, r, noted S = {l → r}. For a word w, S(w) is the set of words obtainable from w by replacing repeatedly l by r. It is clear that, except the particular case l = r, w ∈ S∞ if and only if the set S(w) is infinite. However, despite numerous efforts for more than twenty years, these decidability problems remain open and have become challenging problems. We observe that these problems can be explained in a few minutes to non-scientific people and surely they point out a deep lack of understanding of the rewriting notion. To get new ideas in order to solve these difficult problems, it seems natural to examine closely particular one-rule semi-Thue systems. Several interesting partial results have been obtained in that direction([20, 23, 16, 5, 6, 17]). Here we continue the study of a natural subclass of one-rule semi-Thue systems, introduced by Alfons Geser in [4], and called one-rule grid semi-Thue systems.

It is obvious that the system S = {l → r} with l ≠ r is uniformly terminating if |r| ≤ |l| or if there is a letter x such that |r|x < |l|x. So, when studying the termination, we can assume that there is a letter x such that |r|x > |l|x and there is no letter y such that |r|y < |l|y. Then a borderline case is when there is a unique letter x such that |r|x > |l|x and |r|y = |l|y for y ≠ x. The family of one-rule grid semi-Thue system satisfies a slightly more general condition. This family, that we denote Ŝ † grid, is composed of all one-rule semi-Thue systems having a letter c such that |r|c = |l|c = k > 0.

In [4], Alfons Geser has given a nice decidable characterization of one-rule grid semi-Thue systems that are uniformly terminating by proving that such a system is non-uniformly terminating if and only if it has a loop of length 1 or 2. To know whether a one-rule semi-Thue system has a loop of length 1 or 2 was previously shown decidable by Winfried Kurth in [11].

First, we give a new formulation of this decidable characterization of non-uniformly terminating systems in Ŝ † grid: there exist words x, y such that ly is a left factor of xy and xl is a right factor of ry.

We show that the words x and y are unique and give a simple way to compute these two words. This permits, when considering a nonterminating system S = {l → r} ∈ Ŝ † grid, to give a very precise form of the two words l and r and to get as a consequence that if |r|c = |l|c = k is odd l needs to be a factor of r.

Then we prove that, for any S in Ŝ † grid, the set Loop(S) is a constructible regular language by giving a simple rational expression. This property can not be generalized to arbitrary 1-rule semi-Thue system: for instance Loop( {c → caca} ) is not regular. Concerning the link between Loop(S) and S∞, it is proved in [4] that the family Ŝ † grid satisfies the equality Ŝ † grid = Ŝ † (A∗Loop(S)A∗), that is w ∈ Ŝ † grid if and only if there is a derivation from w to a word in the regular language A∗Loop(S)A∗ where A is the alphabet of lr. Unfortunately, up to now, this relation does not imply that Ŝ † grid is a regular language as shown in [20] when l ∈ a∗b∗. Then we prove that for any S in Ŝ † grid and for any

\footnote{In the original definition of Alfons Geser, a one-rule grid semi-Thue system is a rule semi-Thue system S = {l → r} having a letter c such that |r|c ≤ |l|c. In this paper we do not consider the case |r|c < |l|c for which S is trivially uniformly terminating.}
word $w$, the set $S(w)$ is a constructible context-free language. Note that this result does not hold for arbitrary one-rule semi-Thue systems ([12]). As a matter of fact, $S(w)$ is a bounded context-free language. So we get both the decidability of the termination problem in $S_{grid}$ and the decidability of the common descendant problem in $S_{grid}$. In the last section, we give an argument in favour of the regularity of $S_{\infty}$ by proving that it is a regular language in the case $|r|_c = |l|_c = 2$.

2. Preliminaries and Notations

In the following, $A$ will denote a finite alphabet, $A^*$ the free monoid over $A$ and $e$ the empty word in $A^*$. For a word $w \in A^*$ and a letter $a \in A$, $|w|$ denotes the length of the word $w$ and $|w|_a$ denotes the number of occurrences of the letter $a$ in the word $w$.

Two words $u$ and $v$ are conjugate if there exist words $x$ and $y$ such that $u = xy$ and $v = yx$. It is well known that two words $u$ and $v$ are conjugate if and only if there exists a word $z$ such that $uz = vz$.

A word $u$ is a factor of a word $v$ if there exist two words $w_1$ and $w_2$ such that $v = w_1uw_2$ and we denote by $F(v)$ the set of the factors of the word $v$. We denote by $RF(w)$ (respectively $LF(w)$) the set of right factors (respectively left factors) of the word $w$, that is:

$$RF(w) = \{ v \in A^* \mid \exists u \in A^*, w = uv \},$$

$$LF(w) = \{ u \in A^* \mid \exists v \in A^*, w = uv \}.$$

A semi-Thue system over an alphabet $A$ is a subset $S \subseteq A^* \times A^*$. Members of $S$ are denoted $l \rightarrow r$ (or $l \rightarrow r$ if there is no ambiguity). One-step derivation, denoted $\rightarrow_1$ (resp. $\rightarrow$ if no ambiguity), is the binary relation over words defined by: $\forall u, v \in A^*, u \rightarrow v$ if and only if there exists $l \rightarrow r \in S$ and $\alpha, \beta \in A^*$ such that $u = \alpha l \beta$ and $v = \alpha r \beta$. The relation $\rightarrow_1$ (resp. $\rightarrow$) is the reflexive and transitive closure (resp. transitive closure) of the relation $\rightarrow$ and, for any word $w \in A^*$, we shall denote $S(w)$ the set $S(w) = \{ w' \in A^* \mid w \rightarrow_1 w' \}$ and $S^{-1}(w) = \{ w' \in A^* \mid w' \rightarrow_1 w \}$. We extend these notations to languages: for any language $L \subseteq A^*$, $S(L) = \cup_{w \in L} S(w)$ and $S^{-1}(L) = \cup_{w \in L} S^{-1}(w)$. For a derivation $w = w_0 \rightarrow w_1 \rightarrow \ldots \rightarrow w_n = w'$, $n$ is called the length of the derivation that will be denoted by $w \rightarrow^n w'$.

We note $w \rightarrow_1^n S_{\infty}$ if there is an infinite derivation starting on $w$ and we denote by $S_{\infty}$ the set $S_{\infty} = \{ w \in A^* \mid w \rightarrow_1^n S_{\infty} \}$.

For a semi-Thue system $S$, for any positive integer $n$, we denote $\text{Loop}_n(S) = \{ w \in A^* \mid \exists x, y \in A^*, w \rightarrow^n xwy \}$ and $\text{Loop}(S) = \cup_{n \geq 0} \text{Loop}_n(S)$.

The termination problem for the alphabet $A$ and the semi-Thue system $S \subseteq A^* \times A^*$ is the following:

**instance:** a word $w \in A^*$

---

2This result already appears implicitly inside proofs used by Alfons Geser in [4] to solve the uniform termination problem in $S_{grid}$. 
question: Does every derivation (modulo $S$) starting on $w$ have finite length? (that is does $w \not\in S_\infty$?)

The uniform termination problem for a class $S$ of semi-Thue systems is the following:

instance: an alphabet $A$ and a finite semi-Thue system $S \subseteq A^* \times A^*$ which belongs to $S$

question: Do all derivations (modulo $S$) starting from all words $w \in A^*$ have finite length? (that is does $S_\infty = \emptyset$?)

To get shorter, we say that a system $S$ is nonterminating if the uniform termination problem has a negative answer for $S$. In the sequel, we focus on $S_{\text{grid}}$, the family of one-rule grid semi-Thue systems.

Definition 2.1. A one-rule grid semi-Thue system $S = \{u \rightarrow v\}$ is a one-rule semi-Thue system such that there exists a letter $c$ with $|u|_c = |v|_c = k > 0$.

3. Uniform termination in $S_{\text{grid}}$

In this section, we state some consequences of the following Geser’s result:

Proposition 3.1. ([4]) A one-rule grid semi-Thue system $S = \{u \rightarrow v\}$ is non-uniformly terminating iff one of the following properties is satisfied:

1. $u$ is a factor of $v$.
2. there exist words $g, h, k$ such that $u = gh$, $v = hk$ and $ggh$ is a factor of $hkk$.

Example 3.1. The simplest example of one-rule grid semi-Thue system satisfying the property 2 of proposition 3.1 but not the property 1 is $S = \{cac \rightarrow acca\}$ with $g = c$, $h = ac$, $k = ca$ and $ggh = caccac$ is a factor of $hkk = accaca$.

Let $S = \{u \rightarrow v\} \in S_{\text{grid}}$. The words $u$ and $v$ belong to $A^*$ where $A$ is an alphabet that contains a letter $c$ with $|u|_c = |v|_c = k > 0$. If we denote $B = A \setminus \{c\}$, then

- $u = lc_{u_1 \ldots u_{k-1}c}r$
- $v = l'c_{v_1 \ldots v_{k-1}c}r'$

for some words $l, r, l', r', u_1, \ldots, u_{k-1}, v_1, \ldots, v_{k-1} \in B^*$. We get, as a consequence of the proposition 3.1, the following corollary that is also directly proved in [13]:

Corollary 3.1. Let $S = \{u \rightarrow v\} \in S_{\text{grid}}$ with $u \neq v$. Then $S$ is nonterminating iff one of the two following properties is satisfied:

(i) $u$ is a factor of $v$
(ii) there exist $x, y, s, e \in A^*$ such that $xv = uy e$ and $vy = sxu$.

More precisely, Alfons Geser proved in [4] that a one-rule grid semi-Thue system $S = \{u \rightarrow v\}$ is non-uniformly terminating if and only if $\text{Loop}_1(S) \cup \text{Loop}_2(S) \neq \emptyset$ that is proved decidable by Winfried Kurth in [11].
Proof: Clearly, if (ii) is satisfied, $S$ is nonterminating since $xu \rightarrow xv = uye \rightarrow vye = sxue$. Now, if $u \not\in F(v)$ and if the property 2 of the proposition 3.1 is satisfied, then $hkk = sgghe$ for some $s,e$. Since $|g|_e = |k|_e$, it follows $s,e \in B^+$. Suppose that $g,k \in B^+$ then, since $h \in l'cA^*$ and $hkk = sgghe$, we get $l' = sggl'$ that implies $g = e$ and $u = h \in F(v)$, a contradiction. Thus $|g|_e = |k|_e > 0$ and $k = ye$ for some $y \in A^*$. Taking $x = g$, we get $xv = ghk = uye$, $vye = vk = hkk = sgghe = sxue$, hence $vy = sxu$. □

The following lemma states that if $u \neq v$ the conditions (i) and (ii) of the corollary 3.1 are mutually exclusive. This lemma also clarifies the relationship between the words $x,y,s$ and $e$ when the property (ii) of the corollary 3.1 is satisfied.

Lemma 3.1. Let $S = \{u \rightarrow v\} \in S_{\text{grid}}$. If the property (ii) of the corollary 3.1 is satisfied and $u \neq v$, we have

(i) $|x|_e = |y|_e > 0$ and $s,e \in B^+$,
(ii) $u \not\in F(v)$,
(iii) $u = xu'' = u'y$ where $u''$ is the longest word in $RF(u) \cap LF(v)$ and $u'$ is the longest word in $LF(u) \cap RF(v)$,
(iv) $x$ and $y$ are conjugate; $s$ and $e$ are conjugate.

Proof: If the property (ii) of the corollary 3.1 is satisfied, there are $e$ and $s$ in $A^*$ such that $xv = uye$ (1) and $vy = sxu$ (2). From (1), we get $|x|_e = |ye|_e$ and from (2), $|y|_e = |sx|_e$. So $|x|_e = |y|_e$ and $|e|_e = |s|_e = 0$ that is $e,s \in B^*$. Moreover $|x|_e = |y|_e > 0$ otherwise $x$ and $y$ belong to $B^*$ and, from (1), we obtain $x' = l'$ and from (2), $l' = sxl$. It follows $x = s = e$ and we similarly obtain $y = e = e$. This leads to the contradiction $u = v$. Hence $x \in lcA^*$ and $y \in A^*cre$ so $v$ belongs to $slcA^* \cap A^*cre$ and $l' = sl$, $r' = re$.

Let us now suppose that $u$ is a factor of $v$. We prove that this leads to a contradiction. Indeed, if $v = s'tue'$ for some $e', s' \in B^*$, then the equality (1) gives $xse' = uye \in A^*cre' \cap A^*cre$ and it follows that $e' = e$. Similarly equality (2) implies $s' = s$ and we obtain $xs'u = uy$ and $ue'y = xu$. Therefore $e' = s' = e$ and $u = v$; this contradiction proves (ii).

We can now finish the proof of (i): from (1) and (2), we get $xsxu = xvy = uye$ that implies $u \in LF((xsx)^*)$. Suppose $s = e$ then $u \in LF((x)^*)$ and $vy = sxu \in LF(x^*)$. It follows $v \in LF(x^*)$ that implies $u \in F(v)$, a contradiction. Similarly we can show $e \neq e$.

For (iii), let us first observe that $|x| < |u|$ and $|y| < |u|$; otherwise, we may assume that $|y| \leq |x|$; then $|x| \geq |u|$ and $x = uz$ for some word $z$. It follows that $xu = uz \in RF(vy)$ with $|y| \leq |x| = |zu|$ and $u$ is a factor of $v$ which leads to a contradiction. Then we have $u = xu'' = u'y$ with $u''$ in $RF(u) \cap LF(v)$ and $u'$ in $LF(u) \cap RF(v)$. We shall now prove that $u''$ is the longest word in $RF(u) \cap LF(v)$ and we could similarly show that $u'$ is the longest word in $LF(u) \cap RF(v)$.

We have $x = lc_{u_1} \ldots u_{j-1}u_j'$ with $u_j' \in LF(u_j)$ and, if we denote $p$ the first index such that $u_p \neq u_p$, $p \in \text{F}(v)$, we have $j \leq p$. Assume now that $j < p$ and set $z = lc_{u_1} \ldots u_{p-1}c$, then we have $xsz \in LF(u)$ and $ssz \in LF(v)$. But this implies $s = e$ and $u_p = u_{p-1}c$, a contradiction. Thus $p = j$ and it follows $u_j = u_j'sl$ and $v_j = u_j'l$ with $s \neq e$. Moreover, since $ye \in LF(ue) \cap RF(v)$ with $|x| = |y|$, we obtain $k \geq 2p$ and $xsz' \in LF(u)$, $ssx' \in LF(v)$, with $x' = lc_{u_1} \ldots u_{p-1}c$. 

M. Latteux, Y. Roos / On one-rule grid semi-Thue systems 5
In order to show that $u''$ is the longest word in $RF(u) \cap LF(v)$, assume that there exists some $z$ such that $|z| < |x|$ and $u \in LF(zv)$. Then $|z|_c < |x|_c$, $zsz \in LF(xsx')$ and $sxsx = szsx$ with $t \neq e$. Hence $sx$ is not a primitive word and $xs = (x,s)^{q+1}$ where $x,s$ is the primitive root of $xs$. That implies $x = (x,s)^{q}x_r$ and $x' = (x,s)x'_r$ with $x'_r \in LF(x_r)$, $|x'_r|_c = |x_r|_c > 0$, $z = (x,s)^{r}x_r$ with $i < q$ and $zsz = xsx$. Since $u \in LF(zv)$, we have $xsx' \in LF(zsxz') = LF(xsxxx')$ and $x' = (x,s)^{q}x'_r \in LF(xz') = LF((x,s)^{r}x_r,x,sx'_r)$. Thus $(x,s)^{r}x'_r \in LF(x_r,x,sx'_r)$ and $sx'_r \in LF(x,s)$. But that implies $s = e$, a contradiction that finishes the proof of (ii). Observe that it implies the unicity of $x$ and $y$. Moreover, if there are several $c$ such that $|u|_c = |v|_c$, the value of $x$ and $y$ does not depend on the choice of a particular $c$.

Let us now prove (iv). Since $|x|_c = |y|_c$, we can write $x = x_0cx_1 \ldots cx_t$ and $y = y_0cy_1 \ldots cy_t$ with $0 \leq i \leq t, x_i, y_i \in B^*$. From (1) and (2), we can deduce $xy = sxux = uyey$ and it follows that $yey$ and $sxu$ are conjugate. Since $s, e \in B^*$, it is easy to verify that we must have $y_1cy_0 = x_1sx_0$ and $y_1y_0 = x_0x_0$. Suppose that $|y_i| \geq |x_i|$ (the case $|y_i| \leq |x_i|$ is symmetric) then $y_i = x_iz$ for some $z$ and it follows $y_0 = x_0$. Then we get $y_1cy_0 = x_1sx_0$ and $s$ and $e$ are conjugate. Now from the equality $y_1cy_0 = x_1sx_0$ we obtain $x_1c = x_1c' = x_1sy_0y_0$ that implies $zy = sz = xz$ therefore $x$ and $y$ are conjugate. □

Since the equalities $vy = sxu$ and $xv = uye$ imply $xy = sxux = uyey$, it follows that $u \in LF((sxz)^*)$, $v \in LF((yey)^*)$ and $v \in LF((sxz)^*)$. Then we can state the precise form of $u$ and $v$ for a nonterminating one-rule grid semi-Thue system:

**Proposition 3.2.** $S \in S_{grid}$ is nonterminating iff one of the following three conditions is satisfied

1. $u \in F(v)$
2. $u = (sxz)^{n}x'v$ and $v = (sxz)^{n}sx'$ with $n > 0$ and $x' \in LF(x) \cap LF(szx)$
3. $u = (sxz)^{n}sxz'$ and $v = (sxz)^{n}xsx'c$ with $n \geq 0$, $x = x''x'$ and $x''c = ex'''$.

**Proof:** Let us first observe that these three conditions are sufficient by verifying that for any system $S$ satisfying 1, 2 or 3, $Loop(S) \neq \emptyset$. It is clear for 1 since, in this case, $u \in Loop_1(S)$. If $S$ satisfies 2 then $x(xsz)^{n}x'' \in Loop_2(S)$: indeed, in this case we get $x = x''x'$ and $sxz = x''x''$ for some words $x''$, $x'''$ and $x(xsz)^{n}x'' \rightarrow x(xsz)^{n}sxz = (xsz)^{n}x''x''x'x'' = sxz^k x''x''x'x''$. Finally, if $S$ satisfies 3, $x(xsz)^{n}x'' \rightarrow x(xsz)^{n}x''x'c = (xsz)^{n}x''x''x'x'c \rightarrow (xsz)^{n}x''x''x'x'x''c = (xsz)^{n}x''x''x'x'x''c \rightarrow \cdots$, i.e., there exists $x(xsz)^{n}x'' \rightarrow x(xsz)^{n}x''x'c = sxz^k x''x''x'x'c = sxz^k x''x''x'x'c$, therefore $x(xsz)^{n}x''x' \in Loop_2(S)$. It now remains to prove that the condition (1 or 2 or 3) is necessary.

Suppose that 1 is not satisfied, then we can consider three cases since $u \in LF((xsx)^*)$:

(i) $u = (sxz)^{n}x'v$ with $x = x''x''$.

First we can observe that $v = (sxz)^{n}x'v$ since $v \in LF((sxz)^*)$ and $|v| = |u| + |s|$. Moreover $n > 0$ since $|x|_c < |u|_c$. It remains to prove that $x' \in LF(sx)$. The property (ii) of the lemma 3.1 gives $y \in RF(u)$, therefore $y = x''x''$ since $|x| = |y|$. From the equality $xv = uye$ we obtain $x(xsz)^{n}x'' = (xsz)^{n}x''x''x'c$ and $x' \in LF(sx)$ since $|e| = |s|$. 

(ii) $u = (sxz)^{n}x's' \in S_{grid}$.

Since $v \in LF((sxz)^*)$ and $|v| = |u| + |s|$, it follows $v = (sxz)^{n}sxz'$ with $|x'| = |s'|$. Then $x' \in B^*$ since $|u|_c = |v|_c$ and $s \in B^*$. It follows that $x = x''x''$ for some $x''$. Let us now consider the factorization $x = xclt$ with $l \in B^*$ (we know that $|x|_c > 0$) then we obtain
Proof: We consider the derivation
\[ u = u_1ctx' \quad \text{and} \quad v = v_1ctx' \quad \text{with} \quad |ts'| = |tx'| \quad \text{and} \quad ts', tx' \in B^*, \quad u_1, v_1 \in A^*. \]
This is a contradiction since \( u = u_1cr \) and \( v = v_1cr \) with \(|re| > |r|\). The case \( u = (sxz)^nxsx' \quad \text{with} \quad s' \in LF(s) \) is finally impossible.

(iii) \( u = (sxz)^nxsx' \quad \text{with} \quad x' \in LF(x) \).

Since \( v \in LF((sxz)^\ast) \) and, from the property (iv) of the lemma 3.1, \( |v| = |u| + |c| \), we get
\[ v = (sxz)^nxsx'x'' \quad \text{with} \quad |x''| = |c|. \]
Moreover \( v \in RF((yce)^\ast) \) therefore \( x'' = e \) and \( v = (sxz)^nxsx'e. \)

Now, from the equality \( xv = uy \), we can deduce that \( y = x''x' \) since \( |x| = |y| \).

Let us consider now the equality \( vy = sxu \): we get \( (sxz)^nxsx'vy = sx(xsx)^nxsx' \) and it follows \( x''ey = xsx'. \) Since \( x = x'x'' \) and \( y = x''x' \), we finally obtain \( ex'' = x''s. \)
\[ \square \]

Example 3.2. Let \( S = \{ a^2ca^3b \rightarrow a^4ca^3ca \} \). In this case, \( s = e = a^2 \) and from the property (iii) of the lemma 3.1, \( x = a^2ca \). Then we are in the case 3 of the proposition 3.2 with \( n = 0, x' = a^2c \) and \( x'' = a \).

The condition 2 of the proposition 3.2 implies \( x' \in LF(sx') \) so \( x' \in B^* \). It follows that \( |x'|_c = 2n|x|_c \) is even. Similarly, the condition 3 of the proposition 3.2 implies that \( x'' \in B^* \). Then \( |x''|_c = |x|_c \) and \( |u|_c = 2n|x|_c \) is even. We obtain:

Corollary 3.2. Let \( S = \{ u \rightarrow v \} \) such that \( |u|_c = |v|_c \) is odd. Then \( S \) is nonterminating iff \( u \in F(v) \).

4. The construction of Loop(S) with S in Sgrid

The aim of this section is to prove that for any system \( S = \{ u \rightarrow v \} \) in \( S_{grid} \), the set Loop(S) is a constructible regular language. Clearly, if \( u = v \) then Loop(S) = \( A^*uA^* \) and in the rest of this section, we suppose that \( u \neq v \).

We shall use the following lemma:

Lemma 4.1. Let \( S = \{ u \rightarrow v \} \) in \( S_{grid} \) that is nonterminating with \( u \notin F(v) \). If \( z_0cw_0z'_0 \rightarrow z_4cw_4z'_4 \) with \( z_0, z'_0, z_4, z'_4 \in B^* \) and \( |w|_u = |w|_v \) then \( z_0cw_0z'_0 \in B^*(xu + uy)B^* \).

Proof: We consider the derivation \( z_0cw_0z'_0 \rightarrow z_1cw_1z'_1 \rightarrow z_2cw_2z'_2 \rightarrow z_3cw_3z'_3 \rightarrow z_4cw_4z'_4 \) with \( z_1, z'_1 \in B^* \) and \( |w|_{z_1} > |w|_{z'_1} \). Otherwise we can not apply a derivation step on \( z_1cw_1z'_1 \). Since \( u \notin F(v) \), that implies that the first and the last occurrences of the letter \( c \) can not be both involved in a same step of the derivation. From the proposition 3.2, \( |w|_{z_1} > |w|_{z_1} \).

If the derivation involves the first or the last occurrence of the letter \( c \), then \( |w|_{z_1} > |w|_{z_1} \).

From the hypothesis, we are in the first case and \( |w|_{z_1} = |w|_{z_1} \).

We can also suppose that \( z_0 = l \) and \( z'_0 = r \). Then \( u \in LF(lcw_0cr) \cup RF(lcw_0cr) \) and we consider the case \( lcw_0cr \rightarrow w_0 \rightarrow v_0y_0. \) Since \( w_0 \not\subseteq LF(v) \), we necessarily have \( u \not\subseteq RF(v_0y_0) \). It follows that \( v_0y_0 = sx_1u \rightarrow sx_1v = suy_0e \rightarrow svy_0e = s^2x_2ue. \) We can verify that \( |y_0| = |y_1| \) and \( y_0, y_1 \in RF(u) \). Finally \( y_0 = y_1 \) and we get \( v_0y_0 = sx_1u \) and \( x_1v = uy_0e. \) Now, from the property (ii) of the corollary 3.1 and from the property (iii) of the lemma 3.1, we get \( y_0 = y_1 = x \) that imply \( lcw_0cr = uy \). Observe that if \( z_0 = l \) or \( z'_0 = r \) we still have \( z_0cw_0z'_0 \in B^*uyB^* \).

If \( u \in RF(lcw_0cr) \), then \( z_0cw_0z'_0 \in B^*xuB^* \) that proves the lemma.
\[ \square \]
Let us denote, for any word \( w \in A^*cA^* \), \( \text{int}(w) = (B^*)^{-1}w(B^*)^{-1} \cap cA^* \cap A^*c \).

**Lemma 4.2.** Let \( S \) be nonterminating with \( u \neq v \).

1. If \( w \overset{\gamma}{\rightarrow} w' \) with \( u \in F(v) \) or \( |w|_c > |u|_c \), then \( \text{int}(w') \geq |\text{int}(w)| \).
2. If \( w \overset{\gamma}{\rightarrow} w' \overset{\gamma}{\rightarrow} zuz' \) then \( |\text{int}(w)| = |\text{int}(w')| \).

**Proof:** For the property 1, it is sufficient to prove that \( w \rightarrow w' \) implies \( |\text{int}(w')| \geq |\text{int}(w)| \). That is clear if \( u \in F(v) \) since \( |v| > |u| \) and \( \text{int}(v) = \text{int}(u) \). If \( u \notin F(v) \) then the property (ii) of the corollary 3.1 is satisfied. Since \( |w|_c > |u|_c \), we have to consider three cases:

1. \( w = zuw'' \) and \( w' = zvw'' \) with \( z \in B^* \), \( w'' \in (A^* \setminus B^*) \),
2. \( w = w''uz \) and \( w' = w''vz \) with \( z \in B^* \), \( w'' \in (A^* \setminus B^*) \),
3. \( w = w''uuz'' \) with \( w'' \in (A^* \setminus B^*) \).

Let us consider the case 1. From the property (iv) of the lemma 3.1, \( |\text{int}(u)| = |\text{int}(v)| + |e| \), moreover \( u \in A^*cre \) and \( v \in A^*cre \) so we get \( |\text{int}(w')| = |\text{int}(w)| \). The second case can be proved similarly and for the third case, we clearly obtain \( |\text{int}(w')| > |\text{int}(w)| \) since \( |u| < |e| \). For the property 2, we have \( \text{int}(zuw') = \text{int}(w) \) and \( w \in S_{56} \). So, if \( u \notin F(v), |w|_c > |u|_c \) from the lemma 4.1. Then property 1 yields: \( |\text{int}(w)| \leq |\text{int}(w')| \leq |\text{int}(zuw')| = |\text{int}(w)| \). \( \square \)

**Lemma 4.3.** If the property (ii) of the corollary 3.1 is satisfied and \( u \neq v \) then

1. \( v \notin F(xuB^*) \),
2. \( v \notin F(B^*uy) \),
3. \( xu \notin A^*uA^* \),
4. \( uy \notin A^*uA^* \).

**Proof:** From the property (iii) of the lemma 3.1, \( u = qy \) where \( g \) is the longest word in \( LF(u) \cap RF(v) \). Assume first that \( v \in F(xu) \). Since \( v \in scA^* \) and \( xu \in lcA^* \), \( v \notin LF(xu) \). Then \( u = q'y' \) with \( g' \in LF(u) \cap RF(v) \) and \( |y'| < |x| = |y| \). Thus \( |g| < |g'| \), a contradiction. Assume now that \( v \in F(xuB^*) \). Since \( v \in A^*cre \) and \( u \in A^*cr \), we get \( v \in RF(xu) \) and \( v = w \), a contradiction. Assume now that \( xu \in A^*uA^* \). Consider the equality \( uy = sxu \). If \( u \in LF(xu) \) then \( v = su \), a contradiction. Otherwise, \( u = q'y' \) with \( g' \in LF(u) \cap RF(v) \) and \( |y'| < |y| \). Thus \( |g| < |g'| \), a contradiction. Properties 2 and 4 can be proved similarly. \( \square \)

We obtain as a consequence:

**Corollary 4.1.** If the property (ii) of the corollary 3.1 is satisfied and \( u \neq v \) then \( S^{-1}(B^*(xu + uy)B^*) = B^*(xu + uy)B^* \).
Proof: Let us prove that $S^{-1}(B^*uxB^*) \subseteq B^*(ux + uy)B^*$. It is clearly sufficient to consider a single step of derivation. Let $w_1uw_2 \rightarrow w_1vw_2 = w'_1xuw'_2$ with $w_1, w_2 \in A^*$ and $w'_1, w'_2 \in B^*$. From the lemma 4.3, $v \notin F(uxB^*)$ therefore there exist $t \in B^*$ and $z \in A^*$ such that $w'_1 = w_1t$, $w_2 = zw'_2$ and $txu = vz$. Since $v \in slcA^*$ and $x \in lcA^*$ it follows $t = s$ that implies $z = y$ and $w_1uw_2 = w_1uyw'_2 \in B^*uyB^*$. We can prove similarly that $S^{-1}(B^*uyB^*) \subseteq B^*(ux + uy)B^*$. \qed

Proposition 4.1.

1. If $v = su$ with $s \in B^+$ then $\text{Loop}(S) = RF(s^*)uA^*$.
2. If $v = ue$ with $e \in B^+$ then $\text{Loop}(S) = A^*uLF(e^*)$.
3. If $v = sue$ with $s, e \in B^+$ then $\text{Loop}(S) = RF(s^*)uLF(e^*)$.

Proof: Let $S = \{u \rightarrow su\}$ with $s \in B^+$ and $w \in \text{Loop}(S)$. Then $w = \alpha u\beta \rightarrow \alpha su\beta \rightarrow \cdots \rightarrow \zeta u\beta z' \in B^*$. From the lemma 4.2, we have $|\text{int}(\alpha u\beta)| = |\text{int}(\alpha u\beta)|$ and, since $s \in B^+$, we get $\alpha \in B^*$. Moreover we obtain by induction on the length $n > 0$ of the derivation that $\zeta u\beta z' = \alpha^s u\beta$. It follows $\alpha^s = \zeta x$ and $\alpha \in RF(\alpha^s \times n)$ for any integer $i > 0$. Therefore $\alpha \in RF(s^*)$ and $\text{Loop}(S) \subseteq RF(s^*)uA^*$. Since the converse inclusion is immediate we finally get $\text{Loop}(S) = RF(s^*)uA^*$. The other cases can be proved similarly. \qed

Lemma 4.4. If the property (ii) of the corollary 3.1 is satisfied and $u \neq v$ then for any word $\alpha, \beta \in B^*$

1. $S(\alpha xu\beta) = \{\alpha^s x u e^\beta \mid n \geq 0\} \cup \{\alpha^{s+1} u y e^n \beta \mid n > 0\}$.
2. $S(\alpha u y \beta) = \{\alpha^{s+1} u y e^n \beta \mid n \geq 0\} \cup \{\alpha^s x u e^n \beta \mid n > 0\}$.

Proof: We prove that $S(\alpha xu\beta) = \{\alpha^s x u e^n \beta \mid n \geq 0\} \cup \{\alpha^{s+1} u y e^n \beta \mid n > 0\}$: from the lemma 4.3, $xu \notin A^*uA^+$ and $uy \notin A^*uA^+$ therefore the first steps of a derivation from $w$ are $w = \alpha xu\beta \rightarrow \alpha xv\beta \rightarrow \alpha uy\beta \rightarrow \alpha xy\beta = \alpha x u e \beta$ and the property is proved by induction on the length of the derivation. The second property of this lemma can be proved similarly. \qed

Proposition 4.2. If the property (ii) of the corollary 3.1 is satisfied and $u \neq v$ then

$$\text{Loop}(S) = RF(s^*)(ux + uy)LF(e^*)$$

Proof: Let $w \in \text{Loop}(S)$ then $w \xrightarrow{\alpha} zwz'$ with $z, z' \in B^*$ and $n > 0$. We can suppose $n \geq 4$ and there exists a derivation $w \rightarrow w_1 \rightarrow w_2 \rightarrow w_3 \rightarrow w_4 \xrightarrow{\alpha} zwz'$ with $z, z' \in B^*$. From the lemma 4.2, we get $|\text{int}(w_4)| = |\text{int}(w)|$ and, from the lemma 4.1, we get $w = \alpha xu\beta$ or $w = \alpha uy\beta$ with $\alpha, \beta \in B^*$. Let us suppose that $w = \alpha xu\beta$ then $S(w) = \{\alpha^s x u e^n \beta \mid n \geq 0\} \cup \{\alpha^{s+1} u y e^n \beta \mid n > 0\}$ from the lemma 4.4. Moreover, from the lemma 4.3 we deduce that $xu \neq uy$ since $u \notin LF(xu)$. On the other hand, from the property (iii) of the lemma 3.1, we have $xu \in lcA^*e$ and $uy \in lcA^*e$ then it follows that $\text{int}(uy) \neq \text{int}(xu) = \text{int}(w)$. That implies that if $w = \alpha xu\beta \xrightarrow{\alpha} zwz' = \alpha x u \beta z'$ we must have $\zeta u x u \beta z' = \alpha^s x u e^\beta$ for some $n > 0$ therefore $\alpha \in RF(s^*)$ and $\beta \in LF(e^*)$. The case $w = \alpha uy\beta$ is symmetric and we finally get $\text{Loop}(S) \subseteq RF(s^*)(ux + uy)LF(e^*)$. We can easily verify the converse inclusion since $xu \xrightarrow{\alpha} sxue$ and $uy \xrightarrow{\alpha} suye$. \qed
We directly obtain from the two previous propositions:

**Corollary 4.2.** If \( S \in S_{\text{grid}} \) then \( \text{Loop}(S) \) is a constructible regular language.

The condition \( S \in S_{\text{grid}} \) in the corollary 4.2 is necessary as shown by the following example:

**Example 4.1.** Consider \( S = \{ c \to caca \} \). We shall show that \( \text{Loop}(S) \cap c(ca)^+a^* = L = c\{(ca)^n a^n | n \geq p \geq 0 \} \). Since \( c \rightarrow (ca)^{n-1}ca^{n-1} \), it follows that for \( n > p \geq 0 \), \( c(ca)^n a^n \rightarrow c(ca)^n a^n a^n \rightarrow (ca)^n a^n a^n \in L \). For the reverse inclusion, remark that \( (ca + a)^* \) is \( S \)-closed. So, if \( cwe = c(ca)^n a^p \rightarrow \omega c\epsilon \beta \) for some \( \alpha, \beta \in A^* \), then \( \alpha = \epsilon \) and \( w \rightarrow w\beta \) with \( |\beta| > 0 \). Thus \( \beta = \beta' a^p \), \( (ca)^n \rightarrow (ca)^n a^p \beta' \in \text{LF}(S(c)) \) and \( p < n \) since \( \forall z \in \text{LF}(S(c)), |z| > |z|u \).

### 5. Testing membership in \( S_{\infty} \) with \( S \) in \( S_{\text{grid}} \)

The aim of this section is to show that the termination problem of a system \( S \in S_{\text{grid}} \) is decidable. Recall that this problem is, given \( S \in S_{\text{grid}} \) and a word \( w \in A^* \) to know whether \( w \in S_{\infty} \). Clearly, for a system \( S = \{ u \rightarrow v \} \) with \( u \in F(v) \), \( S_{\infty} = A^* u A^* \) and we conjecture that for any system \( S \in S_{\text{grid}} \) the set \( S_{\infty} \) is a constructible regular language. This is true when \( |u| = |v| = 2 \) as it is proved in the next section but here we must find another way in order to prove that \( S_{\infty} \) is a recursive language for any \( S \in S_{\text{grid}} \). More precisely, we show that for any \( S \in S_{\text{grid}} \) and any word \( w \), the set \( S(w) \) is a constructible bounded context-free language.

**Definition 5.1.** ([9]) A language \( L \subseteq A^* \) is said to be **bounded** if there exist words \( w_1, \ldots, w_n \in A^* \) such that \( L \subseteq w_1^* \ldots w_n^* \).

As a matter of fact, it would be sufficient to prove that for any \( S \in S_{\text{grid}} \) and any word \( w \), the set \( S(w) \) is a constructible context-free language in order to decide whether \( S(w) \) is finite. Nevertheless bounded languages have nice structural properties that will also permit to solve in this section the common descendant problem for any \( S \in S_{\text{grid}} \) that is to decide, given two words \( w \) and \( w' \), whether \( S(w) \cap S(w') \) is not empty. We also observe that the image \( S(w) \) of a word \( w \) by an arbitrary one-rule semi-Thue system \( S \) need not be context-free([12]).

Following Alfons Geser in [4], we first establish that if a derivation is applied to a word long enough, this word can then be factorized so that the derivation applies independently on each factor of the word. Let us denote \( N' = \max\{|cu| \mid i < k \} \) and \( N = \max\{N', |l|, |r| \} \).

**Lemma 5.1.** Let \( w = \z'cm'm''cz'' \) with \( m', m'' \in B^* \) and \( z', z'' \in A^* \). If \( |m'| \geq N \) and \( |m''| \geq N \) then \( S(w) = S(z'cm'm''cz'') \).

**Proof:** Clearly \( S(z'cm'm''cz'') \subseteq S(w) \). Conversely, let \( w' \in S(w) \), we prove by induction on the length \( n \) of the derivation from \( w \) to \( w' \) that \( w' \in S(z'cm'm''cz'') \). The base case \( w = w' \) is immediate. Otherwise \( w \rightarrow w'' \xrightarrow{u} w' \). Since \( |m'm''| > N \), the first step of the derivation applies on an occurrence of \( u \) that can only appear in \( z'cm' \) or in \( m''cz'' \). Suppose it is in \( z'cm' \), then \( z'cm' \rightarrow z_1'cm_1' \) and \( |m_1'| \geq |m'| \) and \( w'' = z_1'cm_1'm''cz'' \). Now, we can apply the induction hypothesis on \( w'' \xrightarrow{u} w' \). The other case can be proved similarly. \( \square \)
We directly deduce:

**Lemma 5.2.** Let $z\omega\epsilon\epsilon\omega z'$ with $z, z' \in B^*$ and \( w \in A^* \). If \( |\epsilon\omega| \geq 2 \times (|w|_c + 2) \times N \) then there exist two constructible words \( w', w'' \in (A^* \setminus B^*) \) such that \( z\epsilon\omega\epsilon\omega z' = w'w'' \) and \( S(z\epsilon\omega\epsilon\omega z') = S(w')S(w'') \).

Then we can state:

**Proposition 5.1.** Let \( S \in S_{\text{grid}} \) satisfying the property (ii) of the corollary 3.1.

1. For any word \( w \in A^* \), the set \( S(w) \) is a bounded context-free language.
2. \( S_{\infty} = S^{-1}(A^*(xu + uy)A^*) = S^{-1}(A^*\text{Loop}(S)A^*) \).

**Proof:** The proof is based on an induction over \( |w|_c \). For the property 2, we have only to prove \( S_{\infty} \subseteq S^{-1}(A^*(xu + uy)A^*) \) since the inclusions \( S_{\infty} \subseteq S^{-1}(A^*\text{Loop}(S)A^*) \subseteq S_{\infty} \) are clear.

- If \( |w|_c < |xu|_c \) then \( S(w) \) is finite (therefore bounded context-free) from the lemma 4.1,
- If \( |w|_c = |xu|_c \) then
  - if \( w \not\in B^*(xu + uy)B^* \) then \( w \not\in S^{-1}(B^*(xu + uy)B^*) \) from the corollary 4.1 and it follows that \( S(w) \) is finite from the lemma 4.1,
  - if \( w \in B^*(xu + uy)B^* \) then, from the lemma 4.4, \( S(w) \) is a (constructible) bounded context-free language and \( w \) is in \( S^{-1}(A^*(xu + uy)A^*) \).
- If \( |w|_c > |xu|_c \), we make a new induction on \( K_w = 2 \times (|w|_c) \times N - |\text{int}(w)| \). If \( K_w \leq 0 \) then it follows from the lemma 5.2 that there exist two constructible words \( w', w'' \in (A^* \setminus B^*) \) such that \( w = w'w'' \) and \( S(w) = S(w')S(w'') \). As \( |w'|_c < |w|_c \) and \( |w''|_c < |w|_c \), we can apply the induction hypothesis and it follows that \( S(w) \) is a constructible bounded context-free language. If \( K_w > 0 \), let us denote \( S_1 = \{ w' \in A^* \mid w' \not\in L \} \) and \( S_2 = \{ w' \in A^* \mid w' \in S \} \). Clearly \( S(w) = S_1 \cup (\bigcup_{w' \in S} S(w')) \). The family of bounded languages is closed by finite union, consequently it remains to prove that for any word \( w' \) in \( S_2 \), \( S(w') \) is a bounded context-free language. Since \( w \not\in B^*(xu + uy)B^* \) it follows from the lemma 4.1 and the lemma 4.2 that, for any word \( w' \in S_2 \), \( |\text{int}(w')| > |\text{int}(w)| \). Then \( K_w < K_w \) and, by the induction hypothesis, \( S(w') \) is a bounded context-free language, moreover if \( w' \in S_{\infty} \) then \( w' \in S^{-1}(A^*(xu + uy)A^*) \) and it follows that \( w \in S^{-1}(A^*(xu + uy)A^*) \) which concludes the proof of the proposition.

When \( u \in F(v) \), a similar proof can be used to show that the property 1 is also satisfied in this case. Then we get as a corollary:

**Corollary 5.1.** The termination problem is decidable for any system \( S \in S_{\text{grid}} \).

Since it is proved in [8] that the non-emptiness of the intersection of two bounded context-free languages is decidable, we also obtain:

**Corollary 5.2.** The common descendant problem is decidable for any system \( S \in S_{\text{grid}} \).
6. The special case $|u|_c = |v|_c = 2$

It has been shown in the section 4 that Loop(S) is regular for $S$ in $S_{\text{grid}}$ and that this result does not hold for arbitrary one-rule semi-Thue system. The status of $S_{\infty}$ is different. Gérard Sénizergues has proved in [20] that $S_{\infty}$ is regular when $u \in a^*b^*$, but the regularity of $S_{\infty}$ is an open problem for arbitrary one-rule semi-Thue system. Note that the regularity of $S_{\infty}$ does not hold for two-rules grid semi-Thue system as shown by the example $S = \{aca \rightarrow c, cc \rightarrow cca\}$ since it is easy to see that $S_{\infty} = A^*cA^*cA^*cA^* \cup \{a^iecA^q \mid p \leq i + q\}$. We conjecture that $S_{\infty}$ is a constructible regular language for any $S$ in $S_{\text{grid}}$. As we are unable to prove this fact, we give an argument in favour by proving it in the simpler case $|u|_c = |v|_c = 2$. The following proof shows also that the structure of $S_{\infty}$ can be involved and that the proof in the general case $S$ in $S_{\text{grid}}$ could be tricky.

Clearly the question arises only when $S$ satisfies the property (ii) of the corollary 3.1 since, when $u \in F(v), S_{\infty} = A^*uA^*$. Then we consider in the following a nonterminating system $S = \{u \rightarrow v\}$ with $u = lceu, v = sceu, xu = uye$ and $vy = sxu$ for $u_1, v_1, s, e \in B^*$ and $x, y \in A^*$. Then there exist words $u_i$ and $u_i'$ such that $u_j = u_i'sl = rued_i$, $v_i = u_i'r = lceu_i$ and $y = u_i'cr$. Let us denote $S_{\infty}^{\min} = S_{\infty} \setminus \{(A S_{\infty} \cup S_{\infty} A)\}$, and for any word $w \in A^*$, $L_w = S^{-1}(A^*w)$ and $S_{\infty}^{-1}(A^*w)$ and $R_w = S^{-1}(wA^*) \setminus (S_{\infty} \cup S_{\infty}^{-1}(wA^*)A)$. Let us suppose that there exists a derivation from $w_0$ to a word in $A^*(xu + uy)A^*$ whose length is strictly less than $n$. It follows that for any $i, q_i = u_{j-1}$ and $p_0 \rightarrow p_i, r_0 \rightarrow r_i$ therefore $p_0 \in S^{-1}(A^*u), r_0 \in S^{-1}(rA^*)$. Since $w_0 \in S_{\infty}^{\min}$, it follows $p_0 \in L_{\text{grid}}, r_0 \in R_{\text{grid}}$ and $w_0 \in L_{\text{grid}}cu_{i+1}$. Let us distinguish two cases for the step $w_j \rightarrow w_{j+1}$:

1. $q_j = u_i'r, r_j = u_i'crz$ and $r_{j+1} = v_i'crz$.

Then $p_0cu_i' \in A^*\text{lcue}_{i+1} = A^*x$ and $w' = p_0cu_i'\text{lcue}_{i}crz \in A^*xuA^*$. Clearly, the length of the derivation $w_0 \rightarrow w_j = p_jcq_jcr_j \rightarrow w' = p_jcq_jcr_j$ is strictly smaller than $n$.

2. $q_j = r_{i+1}u', p_j = z\text{lcue}_{i+1}$ and $r_{j+1} = z\text{scue}_{i}$.

Then $u_i'crA^* = yA^*$ and $w' = p_jcq_jcrn = z\text{lcue}_{i}u_i'crn \in A^*yA^*$. Clearly, the length of the derivation $w_0 \rightarrow w_j = p_jcq_jcr_j \rightarrow w' = p_jcq_jcrn$ is strictly smaller than $n$.

Lemma 6.2. $S^{-1}(A^*u) \setminus A^*cu_{i+1} \subseteq S_{\infty}$

Proof: Let $w \in S^{-1}(A^*u) \setminus A^*cu_{i+1}$. Then there exists a derivation: $w_0 \rightarrow w_1 \rightarrow w_2 \rightarrow \cdots \rightarrow w_n \in A^*u$. Let $j$ be the greatest index such that $w_j \notin A^*cu_{i+1}$. Then $w_j = p_ju_i'cr \in A^*u$, and it follows $w \in S^{-1}(A^*uy) \subseteq S_{\infty}$.
Let us denote \( F = \{ cv_1 \} \cup \{ cu_1 crz \mid v_1 = rez \} \). Remark that \( F \) is finite and \( S^{-1}(A^*l)F \subseteq S^{-1}(A^*xl) \).

**Lemma 6.6.** \( L_{xl} \subseteq L_l F \).

**Proof:** Let \( w \in L_{xl} \). Then \( \rightarrow zxl = zcv_1 \). Let us consider two cases:

1. If \( w = w'cv_1 \), we can deduce that \( w' \rightarrow zl \), so \( w' \in S^{-1}(A^*l) \). Since \( w' = w'cv_1 \notin S_\infty \), it follows \( w' \notin S_\infty \) and since \( w' \notin \text{AS}^{-1}(A^*xl) \), it follows \( w' \notin \text{AS}^{-1}(A^*l) \). Then \( w' \in L_l \)
   and \( w = w'cv_1 \in L_l F \).

2. Else \( w = w'crz \) with \( z \in B^* \) and \( rz \notin v_1 \). Then \( \rightarrow pu_1z \rightarrow pu_1z \rightarrow p'cv_1 \). We show that \( rez = v_1 \). Indeed, if it is not the case, \( psle_1cv_1 \in S^{-1}(A^*u) \setminus A^*cu_1cr \subseteq S_\infty \) which implies \( pu_1 \in S_\infty \) and \( w \in S_\infty \). Since \( w = w'crz \rightarrow pu_1z \), it follows \( w'cr \in S^{-1}(A^*u) \setminus S_\infty \subseteq A^*cu_1cr \), so \( w'crz = w''cr_1uz \rightarrow pu_1z \in A^*xl \). Finally \( w'' \in S^{-1}(A^*l) \), \( w'' \notin S_\infty \) and \( w'' \notin \text{AS}^{-1}(A^*l) \) which implies \( w'' \in L_l \) and \( w \in L_l F \).

Let us denote \( H = \{ z \mid l \in \text{RF}(rez) \setminus \text{RF}(rz) \} \). Remark that \( H \) is finite and \( \text{S}^{-1}(A^*u)H \subseteq \text{S}^{-1}(A^*l) \).

**Lemma 6.4.** \( L_l \subseteq \{ l \} + L_u H \).

**Proof:** Let \( w \in L_l \). If \( w \in A^*l \) then \( w = l \) otherwise \( w = w'crz \) with \( l \notin \text{RF}(rz) \). Then there exists a derivation \( w \rightarrow pu_1z \rightarrow pu_1z \rightarrow p'cv_1 \). Suppose that \( l \notin \text{RF}(rez) \), then \( psle_1cv_1 \in S^{-1}(A^*u) \setminus A^*cu_1cr \) and it follows from the lemma 6.2 that \( pu_1z \), and thus also \( w \), is in \( S_\infty \). This contradiction implies \( l \in \text{RF}(rez) \) therefore \( z \in H \) and \( w'cr \in S^{-1}(A^*u) \). Moreover \( w \notin S_\infty \) and \( w \notin \text{AS}^{-1}(A^*l) \) imply that \( w'cr \notin S_\infty \) and \( w'cr \notin \text{AS}^{-1}(A^*u) \). Finally \( w'cr \in L_u \) and \( w \in L_u H \).

**Lemma 6.5.** \( L_u \subseteq L_l cu_1cr \).

**Proof:** Let \( w \in L_u \subseteq \text{S}^{-1}(A^*u) \setminus S_\infty \subseteq A^*cu_1cr \). Then \( w = w'cu_1cr \rightarrow w''cu_1cr \) with \( w'' \in A^*l \). It follows that \( w'' \in \text{S}^{-1}(A^*l) \). Moreover \( w' \notin S_\infty \) and \( w' \notin \text{AS}^{-1}(A^*l) \) that imply \( w' \in L_l \) and \( w \in L_l cu_1cr \).

Since, clearly, \( \text{S}^{-1}(A^*l)cu_1crH \subseteq \text{S}^{-1}(A^*l) \), it follows from the lemmata 6.4 and 6.5 that \( L_l \subseteq \{ l cu_1cr \}^* \subseteq \text{S}^{-1}(A^*l) \). Then, from the lemma 6.3, we get \( L_{xl} \subseteq \{ l cu_1cr \}^* F \subseteq \text{S}^{-1}(A^*l)F \subseteq \text{S}^{-1}(A^*xl) \).

**Lemma 6.6.** \( L_l \subseteq \{ l (cu_1crH) \}^* \subseteq \text{S}^{-1}(A^*l) \) and \( L_{xl} \subseteq \{ l (cu_1crH) \}^* F \subseteq \text{S}^{-1}(A^*xl) \).

Symmetrically, denoting \( F' = \{ v_1c \} \cup \{ zcu_1c \mid v_1 = zsl \} \) and \( H' = \{ z \mid r \in \text{LF}(zsl) \setminus \text{LF}(zl) \} \), we get:
Lemma 6.7. \( R_r \subseteq (H' lcu_1 c)^* r \subseteq S^{-1}(rA^*) \) and \( R_{yg} \subseteq F' (H' lcu_1 c)^* r \subseteq S^{-1}(ryA^*) \).

Finally we obtain the following regular expression for \( S_\infty \):

**Proposition 6.1.** \( S_\infty = A^* (\{(cu_1 cr H)^* (Fc u_1 c + cu_1 cF') (H' lcu_1 c)^* r A\}^* r A^* \)

**Proof:** From the lemmata 6.1, 6.6 and 6.7, we obtain the following inclusions:

\[
S_\infty = A^* S_{\text{min}} S_{\text{c}} \subseteq A^* (L_{cu_1 c} R_{cg} \cup L_{cu_1 c} R_{yg}) A^* \\
\subseteq A^* (l((cu_1 cr H)^* F cu_1 c (H' lcu_1 c)^* r + l((cu_1 cr H)^* cu_1 c F' (H' lcu_1 c)^* r)) A^* \\
\subseteq A^* (S^{-1}(A^* x l) cu_1 c S^{-1}(r A^*) \cup S^{-1}(A^* y) cu_1 c S^{-1}(r y A^*)) A^* \\
\subseteq A^* (S^{-1}(A^* (x u + y u) A^*)) A^* = S_\infty
\]

\[\square\]

**Example 6.1.** Let \( S = \{a^2 c a^5 c \rightarrow a^4 c a^3 c a^2\} \) that was used in the example 3.2. Here \( s = e = l = a^2 \) and \( r = e \). Recall that \( x = a^2 c a \) and \( y = a^3 c a \). We deduce \( F = ca^3 + ca^3 ca \), \( H = e + a \), \( F' = a^3 c \) and \( H' = \emptyset \). Then

\[ S_\infty = (a + c)^*(a^2 c a^5 c + a^2 c a^3 c + \{(a^3 c + ca^3 c) a^3 c + ca^3 c a^3 c\}|a + c)^* \]

## 7. Conclusion and open questions

This paper deals with the family \( S_{\text{grid}} \) of one-rule semi-Thue systems that are a borderline case for termination. Some of the results obtained here are shown to be false for arbitrary one-rule semi-Thue systems. For instance, it is proved in the section 4 that, for \( S \) in \( S_{\text{grid}} \), \( \text{Loop}(S) \) is a constructible regular language whereas for \( S = \{c \rightarrow caca\} \), \( \text{Loop}(S) \) is not regular. It is also proved that, for \( S \) in \( S_{\text{grid}} \), \( S(w) \) is always a context-free language whereas for \( S = \{ba \rightarrow a^2 ba^2\} \), \( S(b^2 a^2) \) is not context-free [12]. At the contrary, some other results obtained here have been already shown to be true for other particular classes of one-rule semi-Thue systems. So we give new arguments in favour of some conjectures concerning arbitrary one-rule semi-Thue systems. For instance, it is proved that, for \( S \) in \( S_{\text{grid}} \), \( S_\infty = S^{-1}(A^* \text{Loop}(S) A^*) \). This equality appears already in [15] and [23] for other subclasses of one-rule semi-Thue systems. So we can think:

**Conjecture 1.** \( S_\infty = S^{-1}(A^* \text{Loop}(S) A^*) \) for any one-rule semi-Thue system \( S \).

Note that the validity of this conjecture would imply that if \( S \) is a nonterminating one-rule semi-Thue system, then \( \text{Loop}(S) \) is not empty. Moreover, it would give a way to decide whether a word initiates an infinite derivation or not, as mentioned in [7]. For \( S \) in \( S_{\text{grid}} \), it is proved that \( A^* \text{Loop}(S) A^* \), the ideal generated by \( \text{Loop}(S) \), is finitely generated. Does this property hold for arbitrary one-rule Thue systems? It is also proved here that, for \( S = \{u \rightarrow v\} \) with \( |u|_c = |v|_c = 2 \), \( S_\infty \) is a constructible regular language. This result was previously proved in [20] by Géraud Sénizergues for one-rule semi-Thue system with a left hand-side in \( 0^* 1^* \). So we can hope that:

**Conjecture 2.** If \( S \) is a one-rule semi-Thue system, then \( S_\infty \) is a constructible regular language.
This property would imply the decidability of the (uniform) termination problem. However, remark that these two conjectures do not give a way to decide whether or not a word is in $\text{Loop}(S)$.

Another possible extension of our results deals with one-rule grid Thue systems. For $S$ in $\mathcal{S}_{\text{grid}}$, one can consider the Thue system $\hat{S} = S \cup S^{-1}$ and ask: Is it true that $\forall w, \hat{S}(w)$ is a context-free language? Is it true that if $S$ and $S^{-1}$ are terminating then $\forall w, \hat{S}(w)$ is a finite language? Remark that this last question can be also raised when $S$ is an arbitrary one-rule Thue system.

References


