HAL open science

On prefixal one-rule string rewrite systems

Michel Latteux, Yves Roos

▶ To cite this version:

Michel Latteux, Yves Roos. On prefixal one-rule string rewrite systems. Theoretical Computer Science, 2019, 795, pp.240-256. 10.1016/j.tcs.2019.07.004 . hal-02289652

HAL Id: hal-02289652

https://hal.science/hal-02289652v1

Submitted on 11 Oct 2019

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 Prefixal One-Rule String Rewrite Systems

Michel Latteux, Yves Roos*

Univ. Lille, CRIStAL, UMR 9189,
59650 Villeneuve d'Ascq, France

Abstract

Prefixal one-rule string rewrite systems are one-rule string rewrite systems for which the left-hand side of the rule is a prefix of the right-hand side of the rule. String rewrite systems induce a transformation over languages: from a starting word, one can associate all its descendants. We prove, in this work, that the transformation induced by a prefixal one-rule rewrite system always transforms a finite language into a context-free language, a property that is surprisingly not satisfied by arbitrary one-rule rewrite systems. We also give here a decidable characterization of the prefixal one-rule rewrite systems whose induced transformation is a rational transduction.

Keywords: One-rule string rewrite system, Context free language

1. Introduction

String rewrite systems are of primordial interest for computational problems. Mainly, the problems that are investigated for rewrite systems are the accessibility problem, the common descendant problem, the confluence problem, the termination and uniform termination problem. For several years they have been intensively studied and several deep results have been obtained. However some intriguing decidability problems remain open even for very simple rewrite systems[1]. The most known among these problems is certainly the decidability of the termination of one-rule rewrite systems, a question that remains open for more than thirty years.

Other problems consider rewrite systems as transformation operations on languages: given a rewrite system S and a word w, $S^*(w)$ is the set of all the descendants of w in the rewrite system S. Thus S induces a transformation relation S^* over languages and, from there, one can wonder how these transformations on languages can interact with the classical families of formal languages

*Corresponding author
Email addresses: `michel.latteux@univ-lille1.fr` (Michel Latteux), `yves.roos@univ-lille1.fr` (Yves Roos)

[1] See [Der05] for a history of these problems and the attempts to solve them.

([Sén90, Sén95]). In particular, a natural question is the following: given two classes of languages $\mathcal{C}_1$ and $\mathcal{C}_2$, and a family of rewrite systems $\mathcal{F}$, is it true that, for every system S in $\mathcal{F}$ and for every language L in $\mathcal{C}_1$, it holds that $S^*(L)$ is in $\mathcal{C}_2$. If the property is satisfied then rewrite systems in $\mathcal{F}$ are said to be $\mathcal{C}_1/\mathcal{C}_2$ and, in the case when $\mathcal{C}_1 = \mathcal{C}_2$, we rather say that rewrite systems in $\mathcal{F}$ *preserve* $\mathcal{C}_1$.

In this context, some families of rewrite systems have been identified as preserving rational languages ($\mathcal{C}_1 = \mathcal{C}_2 = \mathsf{RAT}$) like *k-period expanding* systems ([Leu08]), *deleting systems* ([HW04]) and *match-bounded* systems ([GHW04]) or preserving context-free languages ($\mathcal{C}_1 = \mathcal{C}_2 = \mathsf{CF}$) like systems *with inhibitor* ([McN01])and *inverse match-bounded* systems ([GHW05]).

Even for very simple rewrite systems, the question whether these systems are $\mathcal{C}_1/\mathcal{C}_2$ for some given classes of languages $\mathcal{C}_1$ and $\mathcal{C}_2$ is not always so easy to answer. One-rule rewrite systems are among the simplest rewriting systems since they are defined by only two words u, v over an alphabet A. Clearly, one-rule rewrite systems do not always preserve regular languages: the simplest example of such a one-rule (length-preserving) rewrite system is the system $S = \{ba \to ab\}$ but it has been proved RAT/CF in [CR90] , in the context of a particular class of rewrite systems called semi-commutations. From this, one could think that one-rule rewrite systems are at least FIN/CF where FIN is the class of finite languages, but, rather surprisingly, it is not the case: for the one-rule rewrite system $S = \{ba \mapsto a^2b^2\}$ it has been proved that $S(b^2a^2)$ is not a context-free language([LR12]). Since then, one-rule *grid* rewrite systems, introduced in [Ges02] have been proved FIN/CF in [LR12].

In this paper, we consider *prefixal* one-rule rewrite systems that are systems in the form $S = \{u \mapsto uf\}$ for some word u and some non empty word f, with respect to the FIN/CF property and to the preserving of RAT. More precisely, after technical preliminaries given in Section 3, we prove that every prefixal one-rule rewrite system S is FIN/CF by defining, for every word w, an appropriate context-free grammar that depends on S and w and that generates $S^*(w)$. Then, in section 5, we give a new example, different from the one given in [LR12], of a one-rule rewrite system that is not FIN/CF and we prove that these two examples are the *simplest* examples of non FIN/CF one-rule rewrite systems. In Section 6, we give a decidable characterization of prefixal one-rule rewrite systems $S = \{u \mapsto uf\}$ for which S^* is a rational relation. In particular we prove that it is the case if and only if $S^*(u)$ is a rational language.

2. Preliminaries and Notations

Let A be a finite alphabet, A^* will denote the free monoid over A and ε the empty word in A^*. For a word $w \in A^*$, $|w|$ denotes the length of the word w and, for any letter $a \in A$, $|w|_a$ denotes the number of occurrences of the letter a in w.

A word w' is a *factor* of a word w if there exist two words w_1 and w_2 such that $w = w_1 w' w_2$ and we denote by $\mathsf{F}(w)$ the set of the factors of the word w. We denote by $\mathsf{RF}(w)$ (respectively $\mathsf{LF}(w)$) the set of *suffixes* (respectively *prefixes*) of the word w, that is:

$$\mathsf{RF}(w) = \{w' \in A^* \mid \exists w'' \in A^*, w = w''w'\},$$

$$\mathsf{LF}(w) = \{w' \in A^* \mid \exists w'' \in A^*,\ w = w'w''\}.$$

For two words w and w', we denote $w \wedge_l w'$ the largest word in $\mathsf{LF}(w) \cap \mathsf{LF}(w')$ and we denote $w \wedge_r w'$ the largest word, in $\mathsf{RF}(w) \cap \mathsf{RF}(w')$.

A word $u \in A^+$ is said to be *primitive* if it is not a power of another word, that is, $u = v^n$ with $v \in A^*$ implies $n = 1$ and $v = u$. The *root* of a word $u \in A^+$ is the unique primitive word ρ such that $u = \rho^n$ for some natural number n. A nonempty set $X \subseteq A^*$ is called a *code* if every equation $u_1 u_2 \dots u_m = v_1 v_2 \dots v_n$ with $u_i, v_j \in X$ for all i and j implies $n = m$ and $u_i = v_i$ for all i.

We denote FIN the family of finite languages, RAT the family of rational languages and CF the family of context-free languages.

A rewrite system over an alphabet A is a subset $S \subseteq A^* \times A^*$. Members of S are denoted $u \mapsto v$. We shall denote S^{-1} the system obtained from the system S by reversing the rules of S, that is $u \mapsto v \in S$ iff $v \mapsto u \in S^{-1}$. One-step derivation, denoted $\rightarrow$, is the binary relation over words defined by : $\forall w, w' \in A^*, w \rightarrow w'$ iff there exists $u \mapsto v \in S$ and $\alpha, \beta \in A^*$ such that $w = \alpha u \beta$ and $w' = \alpha v \beta$. The relation $\xrightarrow{*}$, called *derivation relation*, is the reflexive and transitive closure of the relation $\rightarrow$ and we denote $\xrightarrow{+}$the transitive closure of the relation $\rightarrow$. For a derivation $w = w_0 \rightarrow w_1 \cdots \rightarrow w_n = w'$, n is called the length of the derivation. A rewrite system S induces a transformation over languages: for every word $w \in A^*$, we shall denote $S^*(w)$ the set $S^*(w) = \{w' \in A^* \mid w \xrightarrow{*} w'\}$ and $S^+(w)$ the set $S^+(w) = \{w' \in A^* \mid w \xrightarrow{+} w'\}$; then, for every language $L \subseteq A^*$, $S^*(L) = \bigcup_{w \in L} S^*(w)$ and $S^+(L) = \bigcup_{w \in L} S^+(w)$.

Among the different derivations starting from a word w, some of them play a central role in this paper: a *leftmost* derivation from a word w is, like for leftmost derivations in context-free grammars, a derivation where the rewriting is applied, at each step, to the leftmost possible position in the current word. Symmetrically, we will also consider *rightmost* derivations.

We say that a string rewrite system S is *rational* if the relation $\{(w, w') \mid w' \in S^*(w)\}$ is a rational relation and we say that S is FIN/CF, if for every finite language L, $S^*(L) \in \mathsf{CF}$. If $S = \{u \mapsto v\}$ is a one-rule string rewrite system, we say that S is *prefixal* if $v = uf$ for some non-empty word f; symmetrically S is *suffixal* if $v = fu$ for some word f.

In the following sections, a lot of definitions and claims have symmetric counterparts that are not explicitly stated for the sake of clarity.

3. The languages R_S

Considering an arbitrary string rewrite system S and two words w and w' with $w' \in S^*(w)$, it may happen that, for all derivation from w to w', the right edges of some words that are involved in the derivation are necessarily reached: more precisely, we say that a word w' is *right-complete* with respect to a word w if $w' \in S^+(w)$ and there does not exist some words w_1, w_1' and $w_2 \neq \varepsilon$ such that $w = w_1 w_2$, $w' = w_1' w_2$ and $w_1' \in S^+(w_1)$. We get the following obvious property that will often be used in the forthcoming proofs.

Property 1. *Let S be an arbitrary string rewrite system and w, w' be two words with $w' \in S^+(w)$.*

1. *if w' is right-complete with respect to w then all derivation $w \xrightarrow{+} w'$ can be factorized in $w \xrightarrow{*} zl$ and $zr \xrightarrow{*} w'$ for some word z and some $l \mapsto r \in S$*
2. *there exist words w_2, w_1 and w_1' such that $w = w_1 w_2$, $w' = w_1' w_2$ and w_1' is right-complete with respect to w_1.*

This property motivates the following definition in the case when $S = \{u \mapsto v\}$ is a one-rule rewrite system:

$$R_S = S^{*-1}(A^* u) = \{w \in A^* \ \mid \ \exists w' \in A^*, w \xrightarrow{*} w'u\}.$$

From the above general property, it follows for every one-rule rewrite system S and for every word w that $S^*(w) = S^*(w_1)w_2$ where $w = w_1 w_2$ with w_1 being the longest prefix of w that belongs to R_S.

As for the termination problem (see for instance [Ges03]) and the confluence problem ([Wra90]), a lot of properties of a one-rule string rewrite system $S = \{u \mapsto v\}$, in relation with context-free preserving, depends on the existence of overlaps between the left-hand side u and the right-hand side v of the rewriting rule. For every one-rule rewrite system $S = \{u \mapsto v\}$, we denote:

$$\begin{array}{ll} X = \mathsf{LF}(u) \cap \mathsf{RF}(v) \cap A^+, & Z = \mathsf{RF}(u) \cap \mathsf{LF}(v) \cap A^+, \\ U' = uZ^{-1} = \{u' \in A^* \mid u \in u'Z\}, & U'' = X^{-1}u, \\ V' = vX^{-1} & V'' = Z^{-1}v \text{ as depicted in Figure 1.} \end{array}$$

Note that if $|u| > |v|$ then S terminates so S^* preserves FIN. In the case when $|u| = |v|$, either $u = v$ and S^* is the identity function, or $u \neq v$ and S terminates so S^* preserves FIN. Nevertheless, some non trivial results stated later hold in the case $|u| = |v|$, so, in the rest of the paper, we consider only one-rule string rewrite system $S = \{u \mapsto v\}$ with $|v| \geq |u|$ and $u \neq v$. For such systems, one observe that all the derivations from a word w to a word w' have the same length. Indeed, it is clearly true when $|v| > |u|$ and it has been proved in [LR14] in the case $|u| = |v|$.

In this section, we study how to compute the set R_S. While the definition of this set is very simple, its construction is not so easy. As a matter of fact, this language need not be rational as shown in the following example:

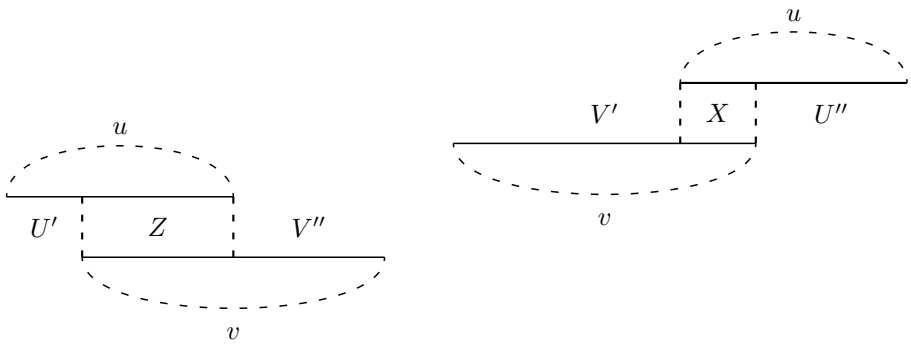


Figure 1: The sets X, Z, U', U'', V' and V''

Example 1. Let $S = \{aab \mapsto aba\}$, then $R_S = A^*Lb$ where $L = \{w \mid |w|_a = |w|_b + 2\}$, so R_S is a context-free language and it is not rational since $R_S \cap a^*b^* = \{a^n b^p \mid n > p\}$. Let us prove the equality $R_S = A^*Lb$.

- $R_S \subseteq A^*Lb$: let $w \in R_S$ and l be the length of a shortest derivation from w to some word of A^*u, say $w'u$. We make an induction on l. If $l = 0$, we get $w \in A^*Lb$ since $A^*aab \subseteq A^*Lb$. If $l > 0$, we get $w = w_1aabw_2 \to w_1abaw_2 \xrightarrow{l-1} w'u$. From the inductive hypothesis, $w_1abaw_2 \in A^*Lb$. We can distinguish four cases that all imply $w_1aabw_2 \in A^*Lb$:
  (i) $w_2 \in A^*Lb$,
  (ii) $aw_2 \in Lb$ that implies $aabw_2 \in Lb$,
  (iii) $baw_2 \in Lb$, that implies $w_2 \in Lb$
  (iv) $w_1'abaw_2 \in Lb$ for some $w_1' \in \mathsf{RF}(w_1)$.

- $A^*Lb \subseteq R_S$: let $w \in L$, we shall prove that $wb \in R_S$ by induction on $|w|$, the length of the word w. One can assume that $w \notin A^+L$. Since $w \in L$, we get $|w| \geq 2$. If $|w| = 2$ then $w = aa$ and $wb \in R_S$. If $|w| > 2$, then $w = a^ibw'$ for some word w' and some integer $i \geq 2$. Since $a^ib \xrightarrow{*} aba^{i-1}$, it follows $w \xrightarrow{*} aba^{i-1}w'$ that implies $a^{i-1}w' \in L$. Now, from the inductive hypothesis, we get $a^{i-1}w'b \in R_S$ so $wb \in R_S$.

We do not know whether there exists some one-rule rewrite system S such that R_S is not a context-free set; we do not even know whether the set R_S is effectively recursive or not. Proposition 3 will give some general bounds for R_S but we can also identify some cases where it is easily seen that this set is rational: for instance, if $u \in \mathsf{RF}(v)$ it clearly follows $R_S = A^*u$ but this condition is not necessary. More precisely, we get:

Proposition 1. *$R_S = A^*u$ if and only if $X \subseteq \mathsf{RF}(u)$.*

Proof. Assume first that $X \not\subseteq \mathsf{RF}(u)$; it follows that there exist some words x, u'' and v' such that $u = xu''$, $v = v'x$ and $x \notin \mathsf{RF}(xu'')$. That implies $u = xu'' \notin \mathsf{RF}(xu''u'')$ and we get $uu'' \notin A^*u$ whereas $uu'' \to vu'' = v'xu'' = v'u = uv''$.

Conversely, assume $X \subseteq \mathsf{RF}(u)$ and consider $w \in R_S$. Then $w \xrightarrow{*} w'u$ for some word w'. We prove that $w \in A^*u$ by induction on the length of the derivation $w \xrightarrow{*} w'u$. If $w = w'u$, we are done else $w = \alpha u\beta \to \alpha v\beta \xrightarrow{*} w'u$ for some words α and β. From the inductive hypothesis, $\alpha v\beta = \alpha' u$ for some word α'. If $|\beta| \geq |u|$ then $w \in A^*u$. Else $v = v'v''$ with $u = v''\beta$ and $v'' \neq \varepsilon$ (recall that $|u| \leq |v|$). So $v'' \in X \subseteq \mathsf{RF}(u)$, and it follows $v''\beta = u \in \mathsf{RF}(u\beta)$ so $w \in A^*u$. □

Observe that Proposition 1 does not hold in the case when $|u| > |v|$: Let $S = \{aab \mapsto a\}$. Since $Z = \emptyset$, we have $Z \subseteq \mathsf{LF}(u)$ but $aaabb \to u$ and $aaabb \notin uA^*$.

Proposition 3 will use the following set in order to give some bounds for R_S:

$$D = \{u'' \in \mathsf{RF}(u) \ \mid \ |vu'' \wedge_r u| > |uu'' \wedge_r u|\}.$$

Why this definition? Consider one step of derivation $w = u'uu'' \to w' = u'vu''$ in a derivation from a word of R_S to a word of A^*u. This step *"goes in the good direction"* if $|w' \wedge_r u| > |w \wedge_r u|$ that gives $|vu'' \wedge_r u| > |uu'' \wedge_r u|$.

We observe that ε is never in D and it is also easily seen that $uD \cap A^*u = \emptyset$.

Proposition 2. *$D = \emptyset$ if and only if $\mathsf{RF}(v) \cap \mathsf{F}(u) \subseteq \mathsf{RF}(u)$.*

Proof. $D \neq \emptyset$ if and only if there exist words $u'', \alpha, \alpha', \alpha'', \beta$ and two distinct letters x and y such that $uu'' = \alpha y\beta u''$, $vu'' = \alpha' x\beta u''$ and $u = \alpha'' x\beta u''$. It is equivalent to the existence of a word $(x\beta)$ in $\mathsf{RF}(v) \cap \mathsf{F}(u)$ which is not in $\mathsf{RF}(u)$ (since $y\beta \in \mathsf{RF}(u)$). □

Let $d = u \wedge_r v$. Clearly, $dD \subseteq A^*d$ that implies $dD^* \subseteq A^*d$. Since $u \neq v$ and $|v| \geq |u|$, there exists some letter $a \in A$ such that $v \in A^*ad$ and we can characterize D as

$$D = \{u' \in A^* \ \mid \ adu' \in \mathsf{RF}(u)\}.$$

Moreover, when $u \notin \mathsf{RF}(v)$, there exists a letter b, distinct from letter a, such that $u \in A^*bd$. In this case, we have $dD \subseteq A^*bd$: indeed let $u'' \in D$, then $du'' \in \mathsf{RF}(u)$ and $|du''| \geq |bd|$ so $bd \in \mathsf{RF}(du'')$. That implies $dD^+ \subseteq A^*bd$.

Lemma 1. *The three following statements are equivalent:*

1. $u \in \mathsf{RF}(v)$
2. $A^*v \subseteq R_S$
3. $A^*v \cap R_S \neq \emptyset$

Proof. Since $v \in A^*ad$, we have $S^*(A^*v) \subseteq S^*(A^*ad) \subseteq A^*ad$. It follows that if $u \notin \mathsf{RF}(v)$ then $S^*(A^*v) \cap A^*u = \emptyset$ so $A^*v \cap R_S = \emptyset$. That proves $3 \Longrightarrow 1$. Conversely, when $u \in \mathsf{RF}(v)$, we get $A^*v \subseteq A^*u \subseteq R_S$ that implies $A^*v \cap R_S \neq \emptyset$. That proves $1 \Longrightarrow 2 \Longrightarrow 3$. □

Using this lemma together with Proposition 1 and Proposition 2, we get the following diagram:

$$\begin{array}{ccccc} A^*v \subseteq R_S & \Longrightarrow & D = \emptyset & \Longrightarrow & R_S = A^*u \\ \Updownarrow & & \Updownarrow & & \Updownarrow \\ u \in \mathsf{RF}(v) & \Longrightarrow & \mathsf{RF}(v) \cap \mathsf{F}(u) \subseteq \mathsf{RF}(u) & \Longrightarrow & X \subseteq \mathsf{RF}(u) \end{array}$$

Conversely, $R_S = A^*u$ does not imply $D = \emptyset$: for $S = \{babb \mapsto baab\}$, we have $X = \{b\} \subseteq \mathsf{RF}(u)$ so $R_S = A^*u$ while $D = \{b\} \neq \emptyset$.

We can now give some bounds for R_S:

Proposition 3. *$A^*uU''^* \subseteq R_S \subseteq A^*uD^* \subseteq A^*d$.*

Proof. For the inclusion $A^*uU''^* \subseteq R_S$ it is clearly sufficient to prove $uU''^* \subseteq R_S$ which is immediate since for all word $u'' \in U''$ there exists some word v' such that $vu'' = v'u$. The last inclusion $A^*uD^* \subseteq A^*d$ directly comes from the inclusion $dD^* \subseteq A^*d$ that we have seen before. It remains to prove $R_S \subseteq A^*uD^*$. Since this inclusion is clearly true if $D = \emptyset$, we may assume $D = \{u' \in A^* \mid adu' \in \mathsf{RF}(u)\}$ with $u \in A^*bd$ and $v \in A^*ad$.

Let w in R_S and consider a derivation of minimal length $w \xrightarrow{*} zu$ with $z \in A^*$. The proof is by induction over the length of the derivation. If $w = zu$, we are done. Else $w = \alpha_0 u\beta_0 \to \alpha_0 v\beta_0 = \alpha_1 u\beta_1 \xrightarrow{*} \alpha_n v\beta_n = zu$ with $\alpha_i \in A^*$ and $\beta_i \in A^+$ for all $i \in [0, n]$. Let k be such that β_k is of minimal length then, for all $i \in [0, n]$ there exists some word β'_i such that $\beta_i = \beta'_i\beta_k$. It follows $w = w'\beta_k$ for some word w' with $w' \xrightarrow{*} \alpha_k u$. From the inductive hypothesis, we get $w' \in A^*uD^*$.

We also have $\alpha_k v \xrightarrow{*} \alpha_n v\beta'_n$ and, since clearly $S^*(A^*v) \subseteq A^*ad$, it follows $\alpha_n v\beta'_n = w''ad$ for some word w''. Moreover $u \notin \mathsf{RF}(\alpha_k u\beta_k)$ since the derivation $w \xrightarrow{*} zu$ is of minimal length; that implies $u \notin \mathsf{RF}(d\beta_k)$. On the other hand, $w''ad\beta_k = zu$ and we get $ad\beta_k \in \mathsf{RF}(u)$ so $\beta_k \in D$ and $w = w'\beta_k \in A^*uD^*$. $\square$

Remark 1. If D is included in U'', Proposition 3 directly implies $R_S = A^*uD^* = A^*uU''^*$. This gives a second case where the language R_S is an effective rational set (as a matter of fact, it generalizes the first case: $R_S = A^*u$ when $u \in \mathsf{RF}(v)$). Nevertheless $R_S = A^*uD^*$ does not imply that D is included in U'' as for $S = \{u \mapsto v\}$ with $u = aab$ and $v = aaba$. In this case, $U'' = \{ab\}$, $D = \{ab, b\}$ and, since $aabab \to aabaab$ and $aabb \to aabab \to aabaab$, we get that $uD \subseteq R_S$ so $R_S = A^*uD^*$. We also observe that, for $S = \{aab \mapsto aba\}$ that was used in Example 1 where $U'' = \{ab\}$ and $D = \{ab, b\}$, the two inclusions $A^*uU''^* \subseteq R_S \subseteq A^*uD^*$ are strict since R_S is not rational in this case. Symmetrically, U'' is not always included in D: if $u \in \mathsf{RF}(v)$, we get $\varepsilon \in U''$ and $D = \emptyset$ so U'' need not be a code, unlike D.

From the above remark, we shall now identify when the inclusions $A^*uU''^* \subseteq R_S$ and $R_S \subseteq A^*uD^*$ are strict or not; for this, we need to study the different possible relations between D and U''. We first observe that $dU'' \subseteq A^*d$ where $d = u \wedge_r v$, that implies $dU''^* \subseteq A^*d$.

Lemma 2. *If $u \in A^*bd$ and $v \in A^*ad$ for some distinct letters a and b and some word d then, for every word $w \in U''$, the four following statements are equivalent:*

1. $w \notin D$
2. $uw \in A^*u$
3. $u \in \mathsf{RF}(dw)$
4. $|w| \geq |u| - |d|$

Proof. Clearly 3 implies directly 2 and 2 implies 1 from Proposition 4. On the other hand, 3 clearly implies 4 and 4 implies 1 since $w \in D$ implies $adw \in \mathsf{RF}(u)$. To finish the proof of the lemma, it is sufficient to prove that 1 implies 3: let $w \in U'' \setminus D$; since $u \in A^*bd$ and $v \in A^*ad$, it follows $u \notin \mathsf{RF}(v)$ that implies $\varepsilon \notin U''$ so $w \neq \varepsilon$. Since $w \in U''$, it follows $vw \in A^*u$. Moreover, since $v \in A^*ad$ and $adw \notin \mathsf{RF}(u)$, we get $|dw| \geq |u|$ so $u \in \mathsf{RF}(dw)$. □

The three following statements have been proved in [LR15] for one-rule length-preserving string rewrite systems. They respectively correspond to Lemma 12, Lemma 9 and Lemma 10 in [LR15]. Although the proofs have been made in a context where $|u| = |v|$, they never use this equality and hold for the more general case $|v| \geq |u|$ since they only depend on the definition of D.

Proposition 4. $uD^* \cap A^+uA^* = \emptyset$.

The two following lemmas assume that $u \in A^*bd$ and $v \in A^*ad$ for some word d and two distinct letters a and b.

Lemma 3. *Let $w \in D^*$ and a word z that satisfy $zw \in D^*$ with $bd \in \mathsf{RF}(bdz)$ then $z \in D^*$.*

Remark 2. From Proposition 4 it follows that for all $w \in A^*uD^*$, there is a unique $w'' \in D^*$ such that $uw'' \in \mathsf{RF}(w)$. Moreover, it follows from Lemma 3 that D is a code and, because D is a code, there is a unique sequence $\alpha_1, \ldots, \alpha_k$ such that $w'' = \alpha_1 \ldots \alpha_k$ and $\alpha_i \in D$ for each $i \in \{1, \ldots, k\}$.

Lemma 4. *If $bu'bd \in \mathsf{RF}(u)$ for some word u', then $A^*uD^* \cap A^*au'bdD^* = \emptyset$.*

This lemma is useful in order to obtain the following relations between U'' and D:

Lemma 5.

1. $uU''^* \subseteq A^*u(D \cap U'')^*$
2. $A^*uU''^* = A^*u(D \cap U'')^*$
3. $A^*uD^* \cap A^*uU''^* = A^*u(D \cap U'')^*$
4. $D^* \cap U''^* = (D \cap U'')^*$
5. $D^* \setminus U''^* = D^*(D \setminus U'')D^*$

Proof. Clearly, 1 implies 2 that implies 3 and 4 implies 5. Let us prove that 3 implies 4. The inclusion $(D \cap U'')^* \subseteq D^* \cap U''^*$ is obvious. Conversely, let $w \in D^* \cap U''^*$. Then $uw \in A^*uD^* \cap A^*uU''^* = A^*u(D \cap U'')^*$ by 3. Proposition 4 yields $uw \cap A^+uA^* = \emptyset$, so $uw \in u(D \cap U'')^*$, and so $w \in (D \cap U'')^*$. It now remains to prove 1 : let $w \in uU''^*$. If $w = u$ then $w \in A^*u(D \cap U'')^*$ otherwise $w = uw_1w_2 \dots w_p$ for some words $w_1 \dots w_p \in U''$ with $p > 0$. If for every $1 \leq i \leq p$ it holds that $w_i \in D$, we are done, else let us consider the biggest i such that $w_i \notin D$. We get $w_{i+1} \dots w_p \in (D \cap U'')^*$ and $uw_1 \dots w_{i-1} \in A^*d$ since $uU''^* \subseteq A^*d$. It follows $uw_1 \dots w_i \in A^*dw_i$ and, from Lemma 2, $uw_1 \dots w_i \in A^*u$ that implies $w \in A^*u(D \cap U'')^*$. □

We shall now prove a key lemma (Lemma 7) that will be particularly useful in the next section. We first need the following corollary of Proposition 4, Lemma 4 and Proposition 3:

Corollary 1. *Let $S = \{u \mapsto v\}$ be a one-rule rewrite system with $|u| \leq |v|$. For all words $w, \alpha, \alpha' \in A^*$ and $x \in D^*$, the two following properties are satisfied:*

1. *if $\alpha u \to w \xrightarrow{*} \alpha' u$ then $w \in A^*u$,*
2. *if $\alpha ux \to w \xrightarrow{*} \alpha' u$ then $w = \alpha vx$ or $w \in A^*ux$.*

Proof. Observe that 1 is a consequence of 2 in the case when $x = \varepsilon$: indeed from 2 we get that if $\alpha u \to w \xrightarrow{*} \alpha' u$ then $w = \alpha v$ or $w \in A^*u$. Since we have seen that $A^*v \cap R_S \neq \emptyset$ if and only if $u \in \mathsf{RF}(v)$, we always have $w \in A^*u$. It remains to prove 2.

Since $\alpha ux \to w$, we have $\alpha ux = w'uw''$ and $w = w'vw''$ for some words w' and w''. From Proposition 4, it follows $|w''| \geq |x|$. We have to consider three cases:

(i) $|w''| = |x|$. In this case, $w = \alpha vx$,
(ii) $|w''| \geq |ux| - |d|$. In this case, $w \in A^*ux$.
(iii) $|x| < |w''| < |ux| - |d|$. In this case, $w'uw'' = w'u''bu'bdx$ for some words u' and u'' with $bu'bd \in \mathsf{RF}(u)$ and $w = w'v''au'bdx$ for some word v''. It follows from Lemma 4 that $w \notin A^*uD^*$, a contradiction since $w \in R_S$ and $R_S \subseteq A^*uD^*$ from Proposition 3.

□

We also need the following lemma:

Lemma 6. $A^*dD^*D^2 \cap A^*u = \emptyset$.

Proof. This is clearly true if $D = \emptyset$. Else, $u \in A^*bd$ and $v \in A^*ad$ for some distinct letters a and b. Moreover, since $dD^* \subseteq A^*d$ it is sufficient to prove $A^*dD^2 \cap A^*u = \emptyset$. Assume $A^*dD^2 \cap A^*u \neq \emptyset$. So, there exist two words $u_1, u_2 \in D$ with $A^*du_1u_2 \cap A^*u \neq \emptyset$. By $u_2 \in D$, we have $adu_2 \in \mathsf{RF}(u)$ so $A^*du_1u_2 \cap A^*u \subseteq A^*du_1u_2 \cap A^*adu_2$. It follows $A^*du_1u_2 \cap A^*adu_2 \neq \emptyset$ that implies $A^*du_1 \cap A^*ad \neq \emptyset$, a contradiction since $adu_1 \in \mathsf{RF}(u)$ and $u \in A^*bd$. $\square$

This allows us to state the following key lemma:

Lemma 7. *For all $w \in A^*$, $x \in D^*$ and $y \in D^+$, if $wuxy \xrightarrow{*} zu$ for some word z then there exists a word $w' \in A^*$ such that:*

1. $wux \xrightarrow{*} w'u$,
2. $w'vy \xrightarrow{*} zu$.

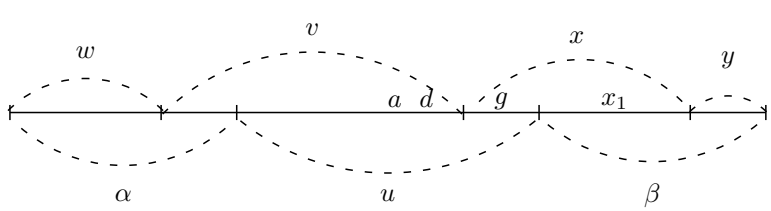
Proof. First observe that the property is clearly true if $D = \emptyset$ by choosing $w' = z$. So we assume $D \neq \emptyset$, $u \in A^*bd$ and $v \in A^*ad$ for some word d and some distinct letters a and b. We prove the property by induction on the length of derivation $wuxy \xrightarrow{*} zu$. Since $y \neq \varepsilon$, it follows from Proposition 4 that the length of the derivation $wuxy \xrightarrow{*} zu$ is at least one.

Assume $wuxy \to zu$. It follows from Corollary 1 that $zu = wvxy$ and, from Lemma 6, $x = \varepsilon$ and $y \in D$. In this case, we can take $w' = w$ that gives the base case of our induction.

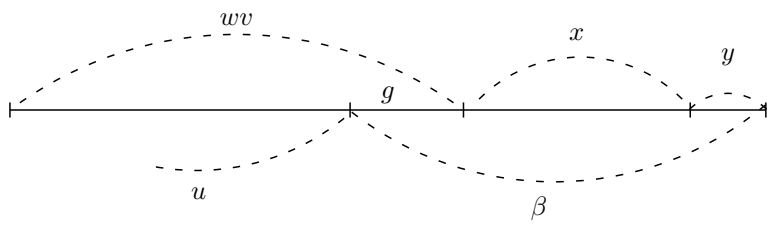
Consider now a derivation $wuxy \to w'' \xrightarrow{*} zu$ for some word w''. From Corollary 1, either $w'' = wvxy$, either $w'' \in A^*uxy$. If $w'' \in A^*uxy$, then $w'' = w'''uxy$ for some word w''' with $wu \xrightarrow{*} w'''u$. From the inductive hypothesis, there exists some w' such that $w'''ux \xrightarrow{*} w'u$ and $w'vxy \xrightarrow{*} zu$. Hence we have $wux \xrightarrow{*} w'''ux \xrightarrow{*} w'u$ and $w'vy \xrightarrow{*} zu$.

It remains the case: $wuxy \to wvxy \xrightarrow{*} zu$. First observe that, if $x = \varepsilon$, we can take $w' = w$ so we assume $x \neq \varepsilon$. Since $wvxy \in R_S$, it follows from Proposition 3 that $wvxy = \alpha u \beta$ for some words $\alpha \in A^*$ and $\beta \in D^*$. Moreover, from Lemma 6, $\beta \neq \varepsilon$. Since $u \notin \mathsf{RF}(v)$ we have $|\beta| < |xy|$ or $|\beta| > |xy|$. We shall distinguish these two cases:

- $|\beta| < |xy|$. From Proposition 4, it follows $u \notin \mathsf{RF}(dD^*)$ so $|u\beta| > |dxy|$. It follows $xy = g\beta$ for some words g with $dg \in \mathsf{RF}(u)$. That implies $adg \in \mathsf{RF}(u)$ so $g \in D$. Since D is a code, it follows $x = gx_1$ with $x_1 \in D^*$ and we get $\alpha u x_1 y \xrightarrow{*} zu$. From the inductive hypothesis, we get that there exists a word w' such that $\alpha u x_1 \xrightarrow{*} w'u$ and $w'vy \xrightarrow{*} zu$. Finally, $wux = wugx_1 \to wvgx_1 = \alpha u x_1 \xrightarrow{*} w'u$.



- $|\beta| > |xy|$. In this case, $\beta = gxy$ for some word g.



That implies $ug \in \mathsf{RF}(A^*v)$ so $d \in \mathsf{RF}(dg)$. On the other hand, we have already seen that $dD^+ \subseteq A^*bd$. Now, from $d \in \mathsf{RF}(dg)$ and $x \in D^+$, we get $dgx \in A^*bd$ which implies $bd \in \mathsf{RF}(bdgx)$. Since $gxy = \beta \in D^*$ and $y \in D^*$, we get $gx \in D^*$ from Lemma 3.

So we can apply the inductives hypothesis on the derivation $wvxy = \alpha ugxy \xrightarrow{*} uz$: there exists a word w' with $\alpha ugx \xrightarrow{*} w'u$ and $w'vy \xrightarrow{*} zu$ that finishes the proof since we get $wux \to wvx = \alpha ugx \xrightarrow{*} w'u$ and $w'vy \xrightarrow{*} zu$.

□

One can observe that it follows from Lemma 7 that $wuxy \in R_S$ with $x, y \in D^*$ implies $wux \in R_S$. We are now able to characterize when $R_S = A^*uU''^*$:

Proposition 5. *$R_S = A^*uU''^*$ if and only if $A^*v(D \setminus U'') \cap A^*uU''^* = \emptyset$.*

Proof. For the *only if* part assume $wvw' \in A^*uU''^*$ for some words $w \in A^*$ and $w' \in D \setminus U''$. Since $wvw' \in A^*uU''^* \subseteq R_S$, it follows that $wuw' \in R_S$. Moreover, since it does not exist $w'' \in D^*$ with $w'' \neq w'$ such that $uw'' \in \mathsf{RF}(wuw')$, it follows that $wuw' \notin A^*u(D \cap U'')^*$ that implies, from Item 2 of Lemma 5, that $wuw' \notin A^*uU''^*$.

Conversely, we assume $R_S \setminus A^*uU''^* \neq \emptyset$, and we shall prove that $A^*v(D \setminus U'') \cap A^*uU''^* \neq \emptyset$ by considering a shortest derivation $w \xrightarrow{*} \beta u$ from some $w \in R_S \setminus A^*uU''^*$ to some word $\beta u \in A^*u$. Since $w \notin A^*u$, there exist words w'', w_0 and w_0' such that $w = w_0uw_0' \to w'' \xrightarrow{*} \beta u$ with $w_0' \in D^* \setminus U''^*$. From the hypothesis on the length of the derivation it follows $w'' \in A^*uU''^*$ so $w'' \notin A^*uw_0'$, and, from Corollary 1, we get $w'' = w_0vw_0'$. On the other hand, it follows $w_0' = xyz$ with $x, z \in D^*$ and $y \in D \setminus U''$ from Item 5 of Lemma 5. Now, by Lemma 7, we get $w_0uxyz \xrightarrow{*} \alpha_1uyz \xrightarrow{*} \alpha_2uz \xrightarrow{*} \beta u$ for some words α_1 and α_2. From the hypothesis on the length of the derivation, it follows $x = z = \varepsilon$ and we get $w'' = w_0vy \in A^*v(D \setminus U'')$. Since $w'' \in A^*uU''^*$, this finishes the proof of the lemma. □

We can also characterize when the equality $R_S = A^*uD^* = A^*uU''^*$ holds:

Proposition 6. *$R_S = A^*uD^* = A^*uU''^*$ if and only if $D \subseteq U''$.*

Proof. We have seen before that if $D \subseteq U''$, the equality $R_S = A^*uD^* = A^*uU''^*$ is a direct consequence of Proposition 3. Conversely, assume $R_S = A^*uD^* = A^*uU''^*$ and let us consider a word $w \in D$. Since $A^*uD^* = A^*uU''^*$, it follows $uw \in A^*uU''^*$. Moreover, from Proposition 4, $uw \notin A^+uA^*$ so $w \in U''^*$. Since $w \in D \cap U''^*$, it follows from Item 4 of Lemma 5 that $w \in D \cap U''$. $\square$

Clearly, $R_S = A^*uD^*$ if and only if for each word $w \in D$, it holds that $uw \in R_S$ and we shall see at the end of this section that this property is decidable. We now state that the inclusions $A^*uU''^* \subseteq R_S$ and $R_S \subseteq A^*uD^*$, given in Proposition 3, can independently be strict or not by considering the four following examples:

(i) As said before, a case where $A^*uU''^* \subsetneq R_S \subsetneq A^*uD^*$ is given in Example 1 where R_S is not rational.

(ii) A case where $A^*uU''^* = R_S = A^*uD^*$ is given by any system S where $u \in \mathsf{RF}(v)$. Indeed, in this case, $D = \emptyset \subseteq U''$ and $R_S = A^*u$.

(iii) A case where $A^*uU''^* = R_S \subsetneq A^*uD^*$ is given by $S = \{u \mapsto v\}$ with $u = aab$ and $v = bba$: indeed we get for this example $D = ab + b$ and $U'' = ab$. Clearly, $ub \notin R_S$ so $R_S \subsetneq A^*uD^*$. On the other hand, $A^*vb \cap A^*u(ab)^* = \emptyset$ and it follows $R_S = A^*uU''^*$ from Proposition 5.

(iv) A case where $A^*uU''^* \subsetneq R_S = A^*uD^*$ is given by $S = \{u \mapsto v\}$ with $u = aab$ and $v = aaba$: in this case, $U'' = ab$ and we have already seen in Remark 1 that $R_S = A^*uD^*$. Moreover $ub \in R_S$ and $ub \notin A^*uU''^*$.

To finish this section, we state that, in the case when $u \in \mathsf{F}(v)$, we obtain a new case where R_S is effective rational set. We need the following result

Proposition 7. *If $u \in \mathsf{F}(v)$ then $\forall w, w' \in A^*, wuw' \in R_S \Longrightarrow uw' \in R_S$.*

Proof. The proof is by induction over the length of a shortest derivation from wuw' to a word of A^*u, say zu. If $wuw' = zu$ then $u \in \mathsf{RF}(uw')$ so $uw' \in R_S$. Else $wuw' = \alpha u\beta$ with $\beta \in D^+$. From Proposition 4, we have $|u\beta| \leq |uw'|$ so $uw' = \gamma u\beta$ for some word γ. Since $\beta \in D^+$, it follows $\alpha u\beta \to w'' \xrightarrow{*} zu$ for some word w''. From Corollary 1 we have to distinguish two cases:

(i) $w'' \in A^*u\beta$: in this case, $w'' = \alpha' u\beta \xrightarrow{*} zu$ for some word α' and it follows from the inductive hypothesis that $u\beta \xrightarrow{*} z'u$ for some word z'. That implies $uw' = \gamma u\beta \xrightarrow{*} \gamma z'u$ so $uw' \in R_S$.

(ii) $w'' = \alpha v\beta$: since $u \in \mathsf{F}(v)$, we have $v = \alpha' u\beta'$ for some words α' and β'. We can apply the inductive hypothesis on the derivation $\alpha\alpha' u\beta'\beta \xrightarrow{*} zu$ and we get $u\beta'\beta \xrightarrow{*} z'u$ for some word z'. It follows $uw' = \gamma u\beta \to \gamma v\beta = \gamma\alpha' u\beta'\beta \xrightarrow{*} \gamma\alpha' z'u$ so $uw' \in R_S$.

$\square$

The construction also needs the set $E = \{x \in D \mid ux \in R_S\}$ that will also be used in the following section. First we prove that E is effectively computable in order to get a decidable property:

Lemma 8. *If $u \in \mathsf{F}(v)$ then the set $E = \{x \in D \mid ux \in R_S\}$ is an effectively computable set.*

Proof. Let $D_1 = \{x \in D \mid vx \in A^*u\}$ and, for all integer $i > 1$, $D_i = \{x \in D \mid vx \in A^*uD_{i-1}^*\} \cup D_{i-1}$. Let $D_\star = \bigcup_{i \geq 1} D_i$. Clearly, since $D_\star \subseteq D$, there exists some integer n such that $D_n = D_{n+1} = D_\star$ so $D_\star$ is an effectively computable set. We shall now prove $D_\star = E$.

> $D_\star \subseteq E$: let $x \in D_\star$ we make an induction on the smallest i such that $x \in D_i$. If $i = 1$, it follows $vx \in A^*u$ so $ux \in R_S$ and $x \in E$. If $i > 1$, it follows $vx \in A^*uD_{i-1}^*$. That implies $vx = wuw'$ for some word $w \in A^*$ and some word $w' \in D_{i-1}^*$. From the inductive hypothesis, we get $w' \in E^*$ so $vx \in R_S$ which implies $ux \in R_S$ and $x \in E$.
>
> $E \subseteq D_\star$: Let $x \in D$ such that $ux \in R_S$. We shall prove $x \in D_\star$ by induction on the length n of a derivation $ux \xrightarrow{n} zu$ with $z \in A^+$. From Proposition 4 we have $ux \to vx \xrightarrow{n-1} zu$ since there is only one occurrence of u in ux and $x \neq \varepsilon$. If $n = 1$, it follows $vx = zu$ so $x \in D_1 \subseteq D_\star$. Else $vx = \alpha u\beta \xrightarrow{n-1} zu$ with $\beta \in D^+$ and, from Proposition 7, $u\beta \in R_S$. Now, from Lemma 7, we get $\alpha u\beta = \alpha u\beta_1 \dots \beta_k \xrightarrow{m} \beta_1' u\beta_2 \dots \beta_k \xrightarrow{p} zu$. Moreover, since all the derivations from a word to another word have the same length, it follows $m + p = n - 1$ so $m < n$ and $p < n$. By induction, we get $\beta_j \in D_\star$ for $1 \geq j \geq k$. Let i be the greatest integer such that $\beta_j \in D_i$ for $1 \geq j \geq k$, we get $vx = u\beta_1 \dots \beta_k \in A^*uD_i^*$ so $x \in D_{i+1} \subseteq D_\star$.

□

And we get as a corollary:

Proposition 8. *If $u \in \mathsf{F}(v)$ then $R_S = A^*uE^*$ so R_S is an effective computable rational set.*

Proof. Since $u \in R_S$ and $uE \subseteq R_S$, we get $A^*uE^* \subseteq R_S$.

Conversely, let $wuw' \in R_S$ with $w' \in D^n$ for some $n \geq 0$. From Proposition 7, it follows $uw' \in R_S$. We shall prove by induction over n that $w' \in E^*$.

- If $n = 0$ then $w' = \varepsilon \in E^*$.
- Else $w' = xy$ with $x \in D^{n-1}$ and $y \in D$. From Lemma 7 and Proposition 7, it follows $ux \in R_S$ and $uy \in R_S$ so $y \in E$ and, from the inductive hypothesis, $x \in E^*$ so $w' = xy \in E^*$ that finishes the proof of the equality $R_S = A^*uE^*$.

□

4. Prefixal one-rule systems are FIN/CF

In this section, we shall prove that the image of all word w by a prefixal one-rule rewrite system $S = \{u \mapsto v\}$ is always context-free by defining a context-free grammar G_w such that $\mathcal{L}(G_w) = S^*(w)$. If $S = \{u \mapsto v\}$ is a prefixal one-rule rewrite system then S clearly satisfies the following property $\mathcal{P}_R$: for all words α and α', if $\alpha u \to \alpha' u$ then $\alpha v \to \alpha' v$. Symmetrically, all suffixal one-rule rewrite system satisfies the property $\mathcal{P}_L$: for all words α and α', if $u\alpha \to u\alpha'$ then $v\alpha \to v\alpha'$. Although property $\mathcal{P}_R$ does not give a characterization of prefixal systems, for all one-rule rewrite system that satisfies it, we shall prove the following nice property: if a word w belongs to $R_S \setminus A^*u$ then a shortest derivation from w to a word in A^*u can be obtained by applying each rewriting step on the rightmost occurence of u. The following lemma shows that properties $\mathcal{P}_R$ and $\mathcal{P}_L$ are decidable.

Lemma 9. *Let $S = \{u \mapsto v\}$ be a one-rule rewrite system with $|u| \leq |v|$. The three following statements are equivalent:*

1. $(\mathsf{RF}(u) \setminus \{u\}) \cap X \subseteq Z$.
2. *S satisfies the $\mathcal{P}_R$ property.*
3. $\forall \alpha, \alpha' \in A^*$, *if* $\alpha u \xrightarrow{*} \alpha' u$ *then* $\alpha v \xrightarrow{*} \alpha' v$.

Proof. $1 \Longrightarrow 2$: Let us consider one step of derivation $\alpha u \to \alpha' u$ for some words α and α'. If $\alpha \to \alpha'$ then $\alpha v \to \alpha' v$. Else $u = u_1 u_2 = u_2 u_3$ and $v = v_1 u_2$ for some words u_1, u_2, u_3, v_1 with $u_2 \neq \varepsilon$ and $\alpha = x u_1$, $\alpha' = x v_1$ for some word x. If $u_1 = \varepsilon$ then $\alpha' u = \alpha v_1 u = \alpha v$ and we get $\alpha v \to \alpha' v$, so we may assume $u_1 \neq \varepsilon$. It follows $u_2 \in (\mathsf{RF}(u) \setminus \{u\}) \cap X$ that implies $u_2 \in \mathsf{LF}(v)$. So we get $v = u_2 v_1'$ for some word v_1' and finally: $\alpha v = x u_1 v_1 u_2 = x u_1 u_2 v_1' \to x v_1 u_2 v_1' = x v_1 v = \alpha' v$.

$2 \Longrightarrow 3$: This is an easy induction, using Corollary 1 that states that if $\alpha u \to w \xrightarrow{*} \alpha' u$ then $w \in A^* u$.

$3 \Longrightarrow 2$: Let us suppose $\alpha u \to \alpha' u$ then $\alpha u \to \alpha' u \to \alpha' v$. From 3, it follows $\alpha v \xrightarrow{*} \alpha' v$ and we get $\alpha u \to \alpha v \xrightarrow{*} \alpha' v$. Since, given two words w and w', all the derivations from w to w' have the same length, it follows $\alpha v \to \alpha' v$.

$2 \Longrightarrow 1$: Let $u_1, u_2, u_3, v' \in A^*$ with $u = u_1 u_2 = u_2 u_3$, $v = v' u_2$ and $u_1, v' \neq \varepsilon$; we have to prove $u_2 \in \mathsf{LF}(v)$. Since $u_1 u = u u_3 \to v u_3 = v' u$, it follows from 2 $u_1 v \to v' v$. One can distinguish two cases:

(i) $u \in \mathsf{LF}(v)$. It follows $u_2 \in \mathsf{LF}(u) \subseteq \mathsf{LF}(v)$.

(ii) $u \notin \mathsf{LF}(v)$. It follows $u \in dbA^*$ and $v \in daA^*$ for some distinct letters a, b and some word d. That implies $u_1 v \in dbA^*$ and $v' v \in daA^*$: indeed, $u_1 u = u u_3 \in dbA^*$ so $u_1 d \in dbA^*$ since $u_1 \neq \varepsilon$ and $v' u = v u_3$ so $v' d \in daA^*$ since $v' \neq \varepsilon$. Since $u_1 v \to v' v$, we get $u_1 v \in uA^*$ that implies $u_2 \in \mathsf{LF}(v)$.

□

The following proposition establishes connections between $\mathcal{P}_R$, $\mathcal{P}_L$, $u \wedge_l v$ and $u \wedge_r v$:

Proposition 9. *Let $S = \{u \mapsto v\}$ be a one-rule rewrite system with $|u| \leq |v|$. Then*

1. *if $|u \wedge_l v| \geq |u \wedge_r v|$ then S satisfies $\mathcal{P}_R$,*
2. *if $|u \wedge_l v| \leq |u \wedge_r v|$ then S satisfies $\mathcal{P}_L$,*
3. *S satisfies $\mathcal{P}_R$ or S satisfies $\mathcal{P}_L$,*
4. *if S is prefixal then S satisfies $\mathcal{P}_R$,*
5. *if S is suffixal then S satisfies $\mathcal{P}_L$,*
6. *if S is confluent then S satisfies both $\mathcal{P}_R$ and $\mathcal{P}_L$.*

Proof. Let $\alpha = u \wedge_l v$ and $\beta = u \wedge_r v$ and assume $|\alpha| \geq |\beta|$. Let us consider $x \in \mathsf{RF}(u) \cap X$, it follows $|x| \leq |\beta|$ that implies $x \in \mathsf{LF}(\alpha) \subseteq \mathsf{LF}(v)$ so $x \in Z$ (a system where $X \cap Z \neq \emptyset$ is called *bordered* in [GZ99]) and S satisfies $\mathcal{P}_R$ from Lemma 9. That proves 1 and symmetrically 2. Items 3, 4 and 5 are direct consequences of 1 and 2. For 6, it has been proved in [Kur90, Wra90] that a one-rule rewrite system $S = \{u \mapsto v\}$ with $|u| \leq |v|$ is confluent if and only if $(\mathsf{LF}(u) \cap \mathsf{RF}(u)) \setminus \{u\} \subseteq \mathsf{LF}(v) \cap \mathsf{RF}(v)$ that implies that S satisfies both $\mathcal{P}_R$ and $\mathcal{P}_L$. □

Observe that, as we said before, property $\mathcal{P}_R$ does not characterize prefixal one-rule systems: $S = \{ba \mapsto aa\}$ satisfies $\mathcal{P}_R$ since $\mathsf{RF}(u) \cap X = \emptyset \subseteq Z$ and S is not prefixal. Nevertheless, we have the following result that will be used later:

Lemma 10. *Let S satisfying $\mathcal{P}_R$. If $\alpha u \beta \xrightarrow{*} \alpha' u$ for some $\alpha, \alpha' \in A^*$ and some $\beta \in D^*$ then the following properties are satisfied:*

1. *if $\beta \neq \varepsilon$ then $\alpha v \beta \xrightarrow{*} \alpha' u$,*
2. *$\alpha v \beta \xrightarrow{*} \alpha' v$,*

Proof. Assume $\beta \in D^+$. From lemma 7, it follows $\alpha u \xrightarrow{*} \alpha'' u$ and $\alpha'' v \beta \xrightarrow{*} \alpha' u$ for some word α''. From Lemma 9, we get $\alpha v \xrightarrow{*} \alpha'' v$ so $\alpha v \beta \xrightarrow{*} \alpha'' v \beta \xrightarrow{*} \alpha' u \rightarrow \alpha' v$. Assume now $\beta = \varepsilon$, we get $\alpha v \xrightarrow{*} \alpha' v$ from Lemma 9. □

This lemma allows us to give the following definition that will be used in the definition of the context-free grammar:

Definition 1. Let S satisfying $\mathcal{P}_R$, we define inductively $\varphi : R_S \mapsto A^* u$ for all word $w \in R_S$ by $\varphi(w) = w$ if $w \in A^* u$ otherwise $\varphi(w) = \varphi(\alpha v \beta)$ if $w = \alpha u \beta$ with $\beta \in D^+$.

This definition is sound since there is no ambiguity for the factorization of arguments of φ from Proposition 4. Moreover, the fact that all the derivations from a word to another have the same length, together with Lemma 10, ensure the termination of this recursive definition. Clearly, we get that if $w \xrightarrow{*} zu$ for some words w and z, it holds that $w \xrightarrow{*} \varphi(w) \xrightarrow{*} zu$. More, thanks to Lemma 7, we can state:

Lemma 11. *If $\alpha u \beta \gamma \xrightarrow{*} \alpha' u$ for some words $\alpha, \alpha' \in A^*$ and $\beta, \gamma \in D^*$, then $\alpha u \beta \gamma \xrightarrow{*} \varphi(\alpha u \beta) \gamma \xrightarrow{*} \varphi(\alpha u \beta \gamma) \xrightarrow{*} \alpha' u$.*

When $u = \varepsilon$, each one-rule rewrite system $S = \{u \mapsto v\}$ is $\mathsf{FIN/CF}$ regardless of the right-hand side v, so, in the rest of this section, we assume that $u \neq \varepsilon$ and that S is prefixal i.e. $v = uf$ for some word $f \neq \varepsilon$ and $u = u_0c$ with $c \in A$. Recall that such a system satisfies $\mathcal{P}_R$.

Proposition 10. *If $wuw' \xrightarrow{*} w''$ for some words $w, w', w'' \in A^*$ then $w'' = w''_1 w''_2$ with $wu_0 \xrightarrow{*} w''_1$ and $uw' \xrightarrow{*} u_0 w''_2$.*

Proof. The proof is an induction on the length of the derivation $wuw' \xrightarrow{*} w''$. The case $wuw' = w''$ is immediate. Else, from Property 1, there exist words w_1, w_2, w'_1, z such that $wuw' = w_1w_2$, $w'_1w_2 = w''$ and $w_1 \xrightarrow{*} zu \to zuf \xrightarrow{*} w'_1$. Let us consider two cases:

(i) $|w_2| > |w'|$. In this case, $w_2 = \gamma c w'$ for some word $\gamma \in A^*$. Let $w''_1 = w'_1\gamma$ and $w''_2 = cw'$. We get $w'' = w''_1 w''_2$, $wu_0 = w_1\gamma \xrightarrow{*} w'_1\gamma = w''_1$ since $w_1 \xrightarrow{*} w'_1$ and $uw' = u_0 w''_2$.

(ii) $|w_2| \leq |w'|$. In this case, $w_1 = wu_0c\gamma$ and $w' = \gamma w_2$ for some word γ with $wu\gamma \xrightarrow{*} zu \to zuf \xrightarrow{*} w'_1$. It follows $wu\gamma = w\gamma' u\gamma''$ for some $\gamma' \in A^*$ and some $\gamma'' \in D^*$. Moreover, $\gamma' u \in uA^*$. From Lemma 10, we get $w\gamma' uf\gamma'' \xrightarrow{*} zuf$ so there exists a word $w_0 = \gamma' uf\gamma''$ with $u\gamma \to w_0$ and $ww_0 \xrightarrow{*} zuf$. Moreover, by $\gamma' u \in uA^*$ follows $w_0 = uw'_0$ for some word w'_0. Since all the derivations from a word to another have the same length, the length of the derivation $wuw'_0 \xrightarrow{*} w'_1$ is strictly smaller than the length of the derivation $wu\gamma \xrightarrow{*} zu \to zuf \xrightarrow{*} w'_1$ so we can apply the inductive hypothesis on the derivation $wuw'_0 \xrightarrow{*} w'_1$ and we get $w'_1 = w''_1 w'_2$ for some words w''_1 and w'_2 with $wu_0 \xrightarrow{*} w''_1$ and $u\gamma \to uw'_0 \xrightarrow{*} u_0 w'_2$. Moreover, denoting $w''_2 = w'_2 w_2$, we have $uw' = u\gamma w_2 \xrightarrow{*} u_0 w'_2 w_2 = u_0 w''_2$.

□

Proposition 10 needs $u \neq \varepsilon$ but one can directly derive the following less precise property that holds in the case $u = \varepsilon$:

Corollary 2. *If $wuw' \xrightarrow{*} w''$ for some words $w, w', w'' \in A^*$ then there exists a factorization $w'' = w''_1 w''_2$ with $wu \xrightarrow{*} w''_1$ and $uw' \xrightarrow{*} uw''_2$.*

We are now able to prove:

Proposition 11. *Each prefixal one-rule rewrite system is* $\mathsf{FIN/CF}$.

Clearly, every prefixal one-rule rewrite system $S = \{\varepsilon \mapsto f\}$ for some word f is $\mathsf{FIN/CF}$ so we assume $u = u_0c$ for some word u_0 and some letter c. In order to prove the proposition, we shall use the following grammar for all prefixal one-rule rewrite system $S = \{u \mapsto uf\}$ and all word w: Let $K_0 = \mathsf{sup}(\{|\varphi(ue)| \mid e \in E\})$ and $K = \mathsf{sup}(\{|w|, K_0, |uf|\})$. Let $A^{\leq K} = \{\alpha \in A^* \mid |\alpha| \leq K\}$ and $B = A^{\leq K} \cup \{\alpha u\beta \mid |\alpha u| \leq K \wedge |u\beta| \leq K\}$. Observe that, clearly, B is closed by factor, that is $B = \mathsf{F}(B)$, and if $ue\beta \in B$ with $e \in E$ and $\beta \in E^*$, then

$\varphi(ue)\beta \in B$: indeed let $e'u = \varphi(ue)$ then $|e'u| \leq K_0 \leq K$ and, from Proposition 4, there is only one occurrence of u in $ue\beta$ so $|ue\beta| \leq K$ and it follows $|u\beta| \leq K$.

We define the grammar $G_w = \langle A, V, S_w, R\rangle$ where

- V is the set of variables, $V = \{S_\alpha \mid \alpha \in B\} \cup \{S'_\alpha \mid \alpha \in B\}$,
- S_w is the axiom,
- and $R \subseteq V \times (V \cup A)^*$ is the set of production rules, $R =$

$$\{S_\alpha \mapsto \alpha, S'_{u_0\alpha} \mapsto \alpha, S_{\alpha\beta} \mapsto S_\alpha S_\beta, S'_{\alpha\beta} \mapsto S'_\alpha S_\beta \mid \alpha, u_o\alpha, \alpha\beta \in B\}$$
$$\cup$$
$$\{S_{\alpha u\beta} \mapsto S_{\alpha u_0} S'_{u\beta}, S'_{\alpha u\beta} \mapsto S'_{\alpha u_0} S'_{u\beta} \mid \alpha u\beta \in B\}$$
$$\cup$$
$$\{S_u \mapsto S_{uf}, S'_u \mapsto S'_{uf}\}$$
$$\cup$$
$$\{S'_{ue\beta} \mapsto S'_{e'u\beta} \mid ue\beta \in B, e \in E, \beta \in E^*, \varphi(ue) = e'u\}$$

Intuitively, a variable S_α may produce $S^*(\alpha)$ and a variable S'_α may produce $u_0^{-1}S^*(\alpha)$. The proof of Proposition 11 consists in the proof of the equality between $S^*(w)$ and $\mathcal{L}(G_w)$. Given a prefixal one-rule rewrite system $S = \{u \mapsto uf\}$ and a word w, we prove the following lemma for the inclusion $S^*(w) \subseteq \mathcal{L}(G_w)$:

Lemma 12. *For all $\alpha \in B$, if $\alpha \xrightarrow[S]{*} \alpha'$ then $S_\alpha \xrightarrow[G_w]{*} \alpha'$ and if $\alpha' = u_0\alpha''$ then $S'_\alpha \xrightarrow[G_w]{*} \alpha''$.*

Proof. The proof is an induction on the length of the derivation in S: $\alpha \xrightarrow{*} \alpha'$. If $\alpha = \alpha'$, we can use the rules $S_\alpha \mapsto \alpha$ and $S'_{u_0\alpha} \mapsto \alpha$, otherwise we may assume that α' is right-complete with respect to α: indeed if $\alpha = \alpha_1\alpha_2$ and $\alpha' = \alpha'_1\alpha_2$ for some words α_1, α'_1 and α_2 with α'_1 right-complete with respect to α_1, there exist the rules $S_{\alpha_1\alpha_2} \mapsto S_{\alpha_1}S_{\alpha_2}$ and $S_{\alpha_2} \mapsto \alpha_2$, moreover, if $\alpha'_1\alpha_2 = u_0\alpha''$ for some word α'', we have $\alpha'_1 = u_0\alpha''_1$ for some word α''_1 since $|\alpha'_1| \geq |u|$ and we can use the rules $S'_{\alpha_1\alpha_2} \mapsto S'_{\alpha_1}S_{\alpha_2}$ and $S_{\alpha_2} \mapsto \alpha_2$.

Since α' is right-complete with respect to α, there exists a derivation $\alpha = \alpha_1 u\alpha_2 \xrightarrow{*} zu \to zuf \xrightarrow{*} \alpha'$ with $\alpha_2 \in E^*$ by Proposition 8. Moreover, from Proposition 10, there exist two words α'_1 and α'_2 such that $\alpha' = \alpha'_1\alpha'_2$ and $\alpha_1 u_0 \xrightarrow{*} \alpha'_1$ and $u\alpha_2 \xrightarrow{*} u_0\alpha'_2$. Note that we cannot have $u\alpha_2 = u_0\alpha'_2$: indeed that would lead to a derivation $\alpha_1 u\alpha_2 = \alpha_1 u_0 c\alpha_2 \xrightarrow{*} \alpha'_1 c\alpha_2 = \alpha'$ that is not right complete, a contradiction. It follows $u\alpha_2 \xrightarrow{+} u_0\alpha'_2$ and we get from the inductive hypothesis on the length of the derivation $\alpha_1 u_0 \xrightarrow{*} \alpha'_1$ that $S_{\alpha_1 u_0} \xrightarrow{*} \alpha'_1$ and, if $\alpha'_1 = u_0\alpha''_1$ then $S'_{\alpha_1 u_0} \xrightarrow{*} \alpha''_1$. It remains to prove $S'_{u\alpha_2} \xrightarrow{*} \alpha'_2$ for which we consider two cases:

(i) $\alpha_2 = \varepsilon$. In this case, $u \to uf \xrightarrow{*} u_0\alpha'_2$. Since $K \geq |uf|$, we get from the inductive hypothesis for the derivation $uf \xrightarrow{*} u_0\alpha'_2$ that $S'_u \to S'_{uf} \xrightarrow{*} \alpha'_2$.

(ii) $\alpha_2 \neq \varepsilon$. There exist some word $\beta \in E^*$ and some word $e \in E$ such that $\alpha_2 = e\beta$. From Proposition 8 and Lemma 11, we get $ue\beta \xrightarrow{+} e'u\beta \xrightarrow{*} u_0\alpha'_2$ with $e'u = \varphi(ue)$. Since $ue\beta \in B$ with $e \in E$, $\beta \in E^*$ and $e'u = \varphi(ue)$, it follows $e'u\beta \in B$ and we get from the inductive hypothesis for the derivation $e'u\beta \xrightarrow{*} u_0\alpha'_2$ that $S'_{e'u\beta} \xrightarrow{*} \alpha'_2$. Moreover $S'_{ue\beta} \mapsto S'_{e'u\beta}$ is a rule in G so $S'_{ue\beta} \to S'_{e'u\beta} \xrightarrow{*} \alpha'_2$.

So we have:

- $S_{\alpha_1 u_0} \xrightarrow{*} \alpha'_1$,
- if $\alpha'_1 = u_0\alpha''_1$ then $S'_{\alpha_1 u_0} \xrightarrow{*} \alpha''_1$,
- $S'_{u\alpha_2} \xrightarrow{*} \alpha'_2$,

and finally we get $S_\alpha = S_{\alpha_1 u\alpha_2} \to S_{\alpha_1 u_0} S'_{u\alpha_2} \xrightarrow{*} \alpha'_1\alpha'_2 = \alpha'$. Moreover, since $|\alpha'_1| \geq |u_0|$, if $\alpha' = u_0\alpha''$ for some word α'', it follows that $\alpha'_1 = u_0\alpha''_1$ for some word α''_1 and we have $S'_\alpha = S'_{\alpha_1 u\alpha_2} \to S'_{\alpha_1 u_0} S'_{u\alpha_2} \xrightarrow{*} \alpha''_1\alpha'_2 = \alpha''$ □

Conversely, an easy induction over the length of the derivations permits to state the following lemma, which gives, together with Lemma 12, a proof for Proposition 11:

Lemma 13. *For all $\alpha \in B$, if $S_\alpha \xrightarrow[G_w]{*} \alpha'$ then $\alpha \xrightarrow[S]{*} \alpha'$ and if $S'_\alpha \xrightarrow[G_w]{*} \alpha'$ then $\alpha \xrightarrow[S]{*} u_0\alpha'$.*

5. Non FIN/CF One-Rule String Rewrite Systems

The aim of this section is to prove that Proposition 11 cannot be generalized to every one-rule string rewrite system and we produce here an example of non FIN/CF one-rule string rewrite system. A example of non FIN/CF system has been provided by Matthias Jantzen in [Jan85, Jan88] in the case of Thue systems i.e. string rewrite systems with symmetric rules. Jantzen proved that for $S = \{\varepsilon \mapsto abba, abba \mapsto \varepsilon\}$ and each word $w \in (a+b)^*$, $S^*(w)$ is not a context-free language. Here, we want to find a one-rule string-rewrite system $S = \{u \mapsto v\}$, as simple as possible, and a word w for which $S^*(w)$ is not a context-free language. The following proposition, that is a consequence of a result of [GHW05], shows that, for such an example, the left-hand side of the rule u must contain at least two distinct letters:

Proposition 12. *Let $S = \{u \mapsto v\}$ be a one-rule string rewrite system. If $u \in a^*$ for some letter a then S is* FIN/CF.

Proof. Let $S = \{a^i \mapsto v\}$ be a one-rule rewrite system over an alphabet A. If $i = 0$ or $v \in a^*$ the property is clearly true. Otherwise there exists some letter b, distinct from a, that occurs in v. Such a system, that is called *rewrite system with an inhibitor* by McNaughton [McN01] have been proved FIN/CF by Geser et al. [GHW05]. □

The example that we produce is the system $S_1 = \{ba \mapsto abab\}$. From above, we cannot find a simplest left-hand side of rule. We shall proceed in two steps: first, find a word w for which $S_1^*(w)$ is not a context-free language and second, prove that we cannot find a simpler example. The proof of the first step uses the fact that the symmetric closure of S_1 is *simplifiable*. This notion of simplifiable system has been already used in [LR14], a previous work on some one-rule string rewrite systems that are equivalent to a rational transduction.

For any rewrite system S, we denote by $S^{\leftrightarrow}$ its symmetric closure: $S^{\leftrightarrow} = S \cup S^{-1} = S \cup \{v \mapsto u \mid u \mapsto v \in S\}$. We say that a rewrite system is *left simplifiable* if for every words w, w' and w'', it holds that $ww'' \in S^*(ww')$ implies $w'' \in S^*(w')$. We define symmetrically the notion of *right simplifiable* rewrite system and we say that a rewrite system is *simplifiable* if it is both left simplifiable and right simplifiable.

Clearly, if a and b are distinct letters then any system $S = \{au \mapsto bv\}$, where u and v are arbitrary words, is left simplifiable. It is also true for $S^{\leftrightarrow}$, the symmetric closure of S as stated in the following lemma. In the following, in order to simplify the notation for the derivations in a symmetric closure of a system $S^{\leftrightarrow}$, we shall write $\leftrightarrow$ instead of $\underset{S^{\leftrightarrow}}{\longrightarrow}$ for a single step of rewriting and $\overset{*}{\leftrightarrow}$ instead of $\overset{*}{\underset{S^{\leftrightarrow}}{\longrightarrow}}$ for a derivation.

Lemma 14. *Let $S = \{u \mapsto v\}$ with $u, v \in A^+$. The three following statements are equivalent:*

1. *S is left simplifiable,*
2. *$S^{\leftrightarrow}$ is left simplifiable,*
3. *$u \in aA^*$ and $v \in bA^*$ for some distinct letters a and b.*

Proof. Clearly 1 is equivalent to 3 and 2 implies 3; it remains to prove 3 implies 2. Assume $u \in aA^*$ and $v \in bA^*$ for some distinct letters a and b, we have to prove that for any words w, w' and w'' it holds that $ww' \overset{*}{\leftrightarrow} ww''$ implies $w' \overset{*}{\leftrightarrow} w''$. The proof is by induction on the length of the derivation $ww' \overset{*}{\leftrightarrow} ww''$. If this length is null, the property is trivially satisfied, otherwise we may clearly assume that $w = x$ for some letter x. If $x \neq a$ and $x \neq b$, the property is clearly true, otherwise we assume $x = a$, the proof in the case $x = b$ being completely symmetric. If no step $\alpha \leftrightarrow \beta$ involved in this derivation satisfies $\alpha = u\gamma$ and $\beta = v\gamma$ then clearly $w' \overset{*}{\leftrightarrow} w''$ so we assume that the derivation $aw' \overset{*}{\leftrightarrow} aw''$ can be factorized in $aw' \overset{*}{\leftrightarrow} u\alpha \leftrightarrow v\alpha \overset{*}{\leftrightarrow} v\beta \leftrightarrow u\beta \overset{*}{\leftrightarrow} aw''$. We can apply the induction hypothesis on the derivation $v\alpha \overset{*}{\leftrightarrow} v\beta$ to get $\alpha \overset{*}{\leftrightarrow} \beta$. This gives a shorter derivation $aw' \overset{*}{\leftrightarrow} u\alpha \overset{*}{\leftrightarrow} u\beta \overset{*}{\leftrightarrow} aw''$ from aw' to aw'' and, by the inductive hypothesis, we get that $w' \overset{*}{\leftrightarrow} w''$. □

Remark 3. Clearly the previous lemma can be symmetrically stated in the case of a system $S = \{ua \mapsto vb\}$ where a and b are two distinct letters. As a consequence, we get that the system $S_1 = \{ba \mapsto abab\}$ is simplifiable.

We can now state:

Proposition 13. *Let $S_1 = \{ba \mapsto abab\}$. The language $S_1^*(baba)$ is not a context-free language.*

Proof. We shall prove that $S_1^*(baba) \cap a^*(ba)^*b^* = \{a^n(ba)^{\mathsf{f}(n)}b^n \mid n \geq 0\}$ where $\mathsf{f}(n)$ is defined inductively by $\mathsf{f}(0) = 2$ and $\mathsf{f}(i+1) = 2 \times \mathsf{f}(i) - 1$. As $(ba)^n \xrightarrow{*} (ab)^{2n} = a(ba)^{2n-1}b$, an easy induction over n proves that $\{a^n(ba)^{\mathsf{f}(n)}b^n \mid n \geq 0\} \subseteq S_1^*(baba)$. Conversely, we observe first that $S_1^*(baba) \cap a^*(ba)^*b^*$ is clearly included in $\{a^n(ba)^pb^n \mid n, p \geq 0\}$. We also observe that for non negative integers i and j, it holds that if $(ba)^i \overset{*}{\leftrightarrow} (ba)^j$ then $i = j$. Indeed assume $j > i$, since S_1 is simplifiable it follows $\varepsilon \overset{*}{\leftrightarrow} (ba)^{j-i}$ a contradiction. This implies, using once again that S_1 is simplifiable, that for every non negative integers n, p and q, if $a^n(ba)^pb^n \overset{*}{\leftrightarrow} a^n(ba)^qb^n$ then $p = q$. So for every non negative integer n, $\mathsf{f}(n)$ is the unique integer p such that $baba \overset{*}{\leftrightarrow} a^n(ba)^pb^n$ and that implies that $S_1^*(baba) = \{a^n(ba)^{\mathsf{f}(n)}b^n \mid n \geq 0\}$ that is not a context-free language. $\square$

We have seen with Proposition 12 that for any example of non $\mathsf{FIN/CF}$ one-rule rewrite system $S = \{u \mapsto v\}$, the left-hand side u must contains two distinct letters so ba is among the two shortest possibilities. In order to show that S_1 is a *smallest* example of non $\mathsf{FIN/CF}$ one-rule rewrite system, we consider now different cases for the right-hand side v.

Let $S = \{ba \mapsto v\}$ be a non $\mathsf{FIN/CF}$ rewrite system. We first observe that both a and b must occur in v otherwise, for every word w, $S^*(w)$ is finite. At last, if $|v|_a = |u|_a$ or if $|v|_b = |u|_b$, such a one-rule string rewrite system S, called grid system in [Ges02] in the context of the uniform termination, has been proved $\mathsf{FIN/CF}$ in [LR12]. All these observations imply that $|v|_a > 1$ and $|v|_b > 1$ so it is not possible to find a smaller example than S_1.

Remark 4.

- For $S_1 = \{u \mapsto v\}$ with $u = ba$ and v=$abab$, clearly, $S_1^*(u) = \{a^n(ba)b^n \mid n \geq 0\}$ and so is a context-free language. It is not always the case: for instance, if we consider the system $S_1' = \{ba \mapsto ababab\}$, a proof very similar to the proof of Proposition 13 permits to show that ${S'}_1^*(w)$ is not context-free as soon as $ba \in \mathsf{F}(w)$.

- In [LR12], another example of a simple non $\mathsf{FIN/CF}$ rewrite system was given: $S_2 = \{u \mapsto v\}$ with $u = ba$ and v=$aabb$ for which $S_2^*(bbaa)$ is not a context-free language. The system S_1 enjoys two properties that are not satisfied by S_2: for S_1, $u \in \mathsf{F}(v)$ and S_1^{-1} is not confluent.

- The systems S_1 and S_2 are the only non $\mathsf{FIN/CF}$ systems in the form $S = \{ba \mapsto v\}$ with $|v|_a = |v|_b = 2$: indeed if $v \in \{abba, baab, baba\}$ then the corresponding system is prefixal or suffixal and for $S_3 = \{u \mapsto v\}$ with $u = ba$ and v=$bbaa$ there exists neither right nor left overlap between u and v nor u itself. That implies that for every word w, the set $S_3^*(w)$ is context-free since it can be obtained by replacing each occurrence of ba in w by the language $\{b^n a^n \mid n > 0\}$. More precisely S_3^* corresponds to the composition of an inverse morphism and a context-free substitution.

6. Decidability of the Rationality

The aim of this section is to give a decidable characterization of prefixal one-rule string rewrite systems that are rational transductions. Observe that this property is stronger than preserving regular languages: for instance every system $S = \{u \mapsto v\}$ with $|v| \leq 1$, called inverse context-free rewrite system in [HW04], preserves regular languages ([BO93]) but need not be rational transductions: $S = \{a\overline{a} \mapsto \varepsilon\}$ corresponds to the *restricted Dyck reduction* for which $(S^{-1})^*(\varepsilon)$ is not a regular language so $(S^{-1})^*$ is not a rational transduction that implies that S^* is not a rational transduction either.

We shall first prove that for any prefixal system $S = \{u \mapsto uf\}$, it holds that $S^*(u) \in \mathsf{RAT}$ if and only if $S^*(u) = uf^*$. Despite the fact that this result holds in both cases $u \in \mathsf{RF}(uf)$ and $u \notin \mathsf{RF}(uf)$, the proofs are different in these two cases. First we state:

Lemma 15. *Let $S = \{u \mapsto uf\}$ be a prefixal one-rule rewrite system such that $u \in \mathsf{RF}(uf)$ and r be the root of f. The three following statements are equivalent:*

1. $S^*(u) \in \mathsf{RAT}$,
2. $S^*(u) = uf^*$,
3. *ur is the shortest member of $A^+u \cap \mathsf{LF}(uf)$.*

Proof. $1 \Longrightarrow 3$: Let $xu = uy$ be the shortest word in $A^+u \cap \mathsf{LF}(uf)$. Hence x and y are primitive words. Let us assume $y \neq r$; it follows $f \notin y^*$ and $f = y^t z$ for some $t > 0$ with $z \neq \varepsilon$ and $y \notin \mathsf{LF}(z)$. We claim that $z \notin \mathsf{LF}(y)$: indeed, since $x^t u = uy^t$, we get $u \in \mathsf{RF}(uy^t)$, and, together with $u \in \mathsf{RF}(uy^t z)$, that implies $u \in \mathsf{RF}(uz)$. If we assume $|z| < |y|$, it follows $uz \notin \mathsf{LF}(uf)$ and $z \notin \mathsf{LF}(y)$. Hence, $\{y, z\}$ is a prefix code. We shall prove the following property, by induction on the length of the derivation:

$$\forall k \geq 0, \forall w \in A^*, x^k uw \in S^*(u) \Longrightarrow k \leq t|w| \tag{1}$$

Let us consider the derivation $u \xrightarrow{*} dug \to dufg = x^k uw$, for some words d and g, with $xu \notin \mathsf{LF}(uw)$ and let us distinguish two cases:

(i) $d = x^p$ for some p. Recall that $x^t u = uy^t$ and $f = y^t z$ so $dufg = dx^t uzg$. Since $xu = uy$ and $\{y, z\}$ is a prefix code, we get $xu \notin \mathsf{LF}(uzg)$. It follows $k = p + t$ and $w = zg$. From the inductive hypothesis, we have $p \leq t|g|$ so $k = p + t \leq t|zg|$.

(ii) $d = x^p x'$ for some p with $x' \neq \varepsilon$ and $x \notin \mathsf{LF}(x')$. We claim that $x'u \notin \mathsf{LF}(x^+u)$: else we have $x'u \in A^+u \cap \mathsf{LF}(uy) \subseteq A^+u \cap \mathsf{LF}(uf)$ and $|x'| < |x|$ which is a contradiction to the definition of x. It follows $k \leq p$ and $dug = x^k uw'$ with $|w'| < |w|$. From the induction hypothesis, we get $k \leq t|w'|$ so $k \leq t|w|$.

Now, since $uf = xuy^{t-1}z$, it follows $\{x^n u(y^{t-1}z)^n \mid n > 0\} \subseteq S^*(u)$ and, from Equation 1, $S^*(u) \notin \mathsf{RAT}$.

$3 \Longrightarrow 2$: Let $ur = xu$ for some x; we claim that if $dug \in uf^*$ for some words d and g then $d \in x^*$. Indeed, if $d = \varepsilon$, it is clear; else if $dug \in uf^*$ then $dug \in ur^* = x^*u = x^*ur^*$ and it follows $d \in \mathsf{LF}(x^*)$. So $d = x^m x'$ for some m and some non empty word $x' \in \mathsf{LF}(x)$. Since $dug \in ur^* = x^*ur^*$, we get $x'u \in \mathsf{LF}(uf)$ and that implies $x' = x$ because xu is the shortest word in $A^+u \cap \mathsf{LF}(uf)$. It follows $d \in x^*$, $g \in r^*$ and $uf^* = S^*(uf^*)$. Since the implication $2 \Longrightarrow 1$ is clearly true, this finishes the proof of the lemma. □

We observe that in the case when $u = \varepsilon$, Lemma 15 states that $S^*(\varepsilon)$ is rational if and only if the root of f is its first letter, so we get the well known result: for all rewrite system $S = \{\varepsilon \mapsto f\}$, $S^*(\varepsilon) \in \mathsf{RAT}$ if and only if $f \in a^*$ for some letter a. In the case when $S = \{u \mapsto v\}$ is a one-rule string rewrite system that satisfies $u \notin \mathsf{RF}(v)$, we can state:

Lemma 16. *Let $S = \{u \mapsto v\}$ be a one-rule string rewrite system with $u \in \mathsf{F}(v) \setminus \mathsf{RF}(v)$. The three following statements are equivalent:*

1. $S^*(u) \in \mathsf{RAT}$,
2. $S^*(u) = uf^*$ *for some word* f,
3. $S^*(u) \cap A^+uA^* = \emptyset$.

Proof. Assume $S^*(u) \cap A^+uA^* = \emptyset$. It follows $v = uf$ for some f and $S^*(u) = uf^*$ that implies $S^*(u) \in \mathsf{RAT}$. It remains to prove that if $S^*(u) \in \mathsf{RAT}$ then $S^*(u) \cap A^+uA^* = \emptyset$. We shall prove the contrapositive: assume $S^*(u) \cap A^+uA^* \neq \emptyset$ and let w_0 be the smallest word of $S^*(u) \cap A^+uA^*$. One can set $w_0 = xuy$ with $x \neq \varepsilon$ and $uy \notin A^+uA^*$. Since $u \notin \mathsf{RF}(v)$, Lemma 1 implies that $S^*(u) \cap A^+u = \emptyset$. Thus $y \neq \varepsilon$ and we can write $y = r^t$ for some primitive word r, and some $t > 0$. If $u \in \mathsf{RF}(r^*)$, it follows $uy \in \mathsf{RF}(r^*)$ and $u \in \mathsf{RF}(uy)$, a contradiction. Thus $u \notin \mathsf{RF}(r^*)$ and u admits a factorization $u = zr^p$ with $p \geq 0$ and $\{z, r\}$ a suffix code. We shall first prove:

$$v = z'r^{p+t} \text{ for some word } z' \text{ with } \{z', r\} \text{ a suffix code and } |z'| > |z| \quad (2)$$

If $w_0 = v$ then we can set $z' = xz$ else $v = uf$ and $w_0 = uf^i$ for some $i > 1$. By definition, $y \in \mathsf{RF}(f^i)$. By minimality of w_0, we have $|y| < |f|$, and so $f = f'y$ for some $f' \in A^+$. We claim that:

$$rf' = z''r^p \text{ for some word } z'' \text{ with } \{z'', r\} \text{ a suffix code} \quad (3)$$

To prove 3, observe first that $f' \notin r^*$: indeed, if $f' \in r^*$ then $uf^i \in ur^*$ that implies $xu \in ur^*$ and $u \in \mathsf{RF}(r^*)$, a contradiction. Now, since r is a primitive word, $rf' \notin \mathsf{RF}(r^*)$ and $rf' = z''r^j$ for some j with $\{z'', r\}$ a suffix code. Moreover $xu = uf^{i-1}f' = xzr^p \in A^*rf' = A^*z''r^j$, that implies $p = j$, so $rf' = z''r^p$ with $\{z'', r\}$ a suffix code. We also have:

$$r \in \mathsf{RF}(uf') \Longleftrightarrow r \in \mathsf{RF}(u) \tag{4}$$

Indeed from $w_0 = xur^t = uf^{i-2}f^2 = uf^{i-2}f'r^tf'r^t$, it follows $xu \in A^*rf'$ so $xuf' \in A^*rf'f'$. Now if $r \in \mathsf{RF}(uf')$, we get $r \in \mathsf{RF}(rf'f')$ so $r \in \mathsf{RF}(rf')$ that implies $xu \in A^*r$. Moreover, $u \notin \mathsf{RF}(r)$ and it follows $r \in \mathsf{RF}(u)$; the converse implication is clear.

We are now able to prove 2 by considering two cases:

(i) $p > 0$. From the equality $v = zr^pf'r^t$, we get, from 3, $v = zr^{p-1}z''r^{p+t}$ and we can set $v = z'r^{p+t}$ with $z' = zr^{p-1}z''$.

(ii) $p = 0$. From 4 it follows $r \notin \mathsf{RF}(uf')$. Moreover, since $uy \notin A^+uA^*$, we get $uf' \notin \mathsf{RF}(r)$ and we can set $v = z'r^{p+t}$ with $z' = uf' = zf'$.

We now use 2 to prove the key equation:

$$\forall w, k, wr^k \in S^*(u) \Longrightarrow k \leq (p+t)|w| \tag{5}$$

The proof is an induction over the length of the derivation from u to wr^k. If $u = wr^k$ then $w = zr^{p-k}$ so $k \leq p+t$ and $|w| \geq |z| > 0$ that implies $k \leq (p+t)|w|$. Else, $u \xrightarrow{*} \alpha u\beta \to \alpha v\beta = wr^k$. Let $\beta = \beta' r^j$ with $r \notin \mathsf{RF}(\beta')$ and let us consider two cases:

(i) if $\beta' = \varepsilon$, we get $p + j \leq (p+t)|\alpha z|$ from the inductive hypothesis. Thus $k \leq p + t + j \leq p|\alpha z| + t(|\alpha z| + 1)$. Since $|z'| > |z|$, it follows $k \leq (p+t)|\alpha z'| \leq (p+t)|w|$.

(ii) if $\beta' \neq \varepsilon$, since r is a primitive word, it follows $r^2 \notin A^*r\beta'$. Thus $k \leq j+1$ and $|wr| \geq |\alpha v\beta'| > |\alpha u\beta'| + |r|$. From the inductive hypothesis, we get $j \leq (p+t)|\alpha u\beta'|$ that implies $j < (p+t)|w|$ and $k \leq j+1 \leq (p+t)|w|$.

Finally, since $xuy \in S^*(u)$, we get $\{x^nuy^n \mid n > 0\} = \{x^nur^{tn} \mid n > 0\} \subseteq S^*(u)$ and it follows from 5 that $S^*(u) \notin \mathsf{RAT}$. □

As a consequence of the two previous lemmas, we get:

Proposition 14. *If $S = \{u \mapsto uf\}$ is rational then $S^*(u) = uf^*$.*

The rest of this section is devoted to the converse of Proposition 14: if $S = \{u \mapsto uf\}$ with $S^*(u) = uf^*$, then S is rational. First, we prove:

Lemma 17. *Let $S = \{u \mapsto uf\}$ be a prefixal one-rule rewrite system. If $S^*(u) = uf^*$ then for all $w \in A^*$, $S^*(uw)$ is a rational language.*

Proof. If $u = \varepsilon$, then $S^*(uw) = f^*t(w)$ where t is the rational substitution defined by: $\forall x \in A, t(x) = xf^*$. Else, $u = u_0c$ for some word u_0 and some letter c. We make an induction on $|w|$, the length of the word w. If $w = \varepsilon$, $S^*(uw) = uf^*$ else, we observe that if $uf^*w \cap A^+uA^* = \emptyset$, it follows $S^*(uw) = uf^*w$ so we assume $uf^*w \cap A^+uA^* \neq \emptyset$ and we denote i, the smallest integer such that $uf^iw \in A^+uA^*$. Let us consider the rightmost occurrence of u in uf^iw: we set $uf^iw = xuy$ with $x \neq \varepsilon$ and $y \notin A^*uA^*$. From Proposition 10, we get $S^*(xuy) = S^*(xu_0)[(u_0)^{-1}S^*(uy)]$. Let us distinguish two cases:

(i) $xu = uf^iw'$ with $w = w'y$ and $w' = w''c$. From the inductive hypothesis, $S^*(uw'')$ and $S^*(uy)$ are rational languages. It follows from the definition of the integer i that $S^*(uf^iw'')$ is a rational language; indeed we get $S^*(uw'') = uw'' + ufw'' + \cdots + uf^{i-1}w'' + S^*(uf^iw'')$ so $S^*(uw'') \setminus S^*(uf^iw'')$ is finite. As a consequence, we get that $S^*(uf^iw) = S^*(uf^iw'')[u_0^{-1}S^*(uy)]$ is rational since . That finally implies that $S^*(uw)$ is a rational language since $S^*(uw) = uw+ufw+\cdots+uf^{i-1}w+S^*(uf^iw)$.

(ii) $xu \in \mathsf{LF}(uf^i)$. In this case, since $S^*(u) \cap A^+uA^* \neq \emptyset$, Lemma 16 implies $u \in \mathsf{RF}(uf)$ that also implies $uw \notin A^+uA^*$. It follows $uf^* \subseteq A^*u$ so $S^*(uw) = uf^*w$ is rational.

□

In the following, we consider a prefixal one-rule rewrite system $S = \{u \mapsto uf\}$ over an alphabet A and we assume $S^*(u) = uf^*$. The system S is rational if and only if the language $M = \{(w, w') \mid w' \in S^*(w)\}$ is a rational subset of $A^* \times A^*$(See [Ber79] for a reference study on rational transductions and their use for the classification of context-free languages.). If we denote $L = \{(w, w') \mid uw' \in S^*(uw)\}$, we observe that if L is rational, so is M: indeed M can be obtained as $M = \mathsf{Id} \cup \mathsf{Id}(u, u)L$ where $\mathsf{Id} = \{(w, w) \mid w \in A^*\}$. So, it remains to prove that L is a rational language.

Let $R = \{w \mid uw \in A^+u \wedge uw \notin A^+uA^+\}$ and $K = \{w \mid uw \notin A^+uA^*\}$. Intuitively, R permits to point out all the left factors of uw that are in A^+u. Clearly $\varepsilon \in K$ and $A^* = R^*K$. One can also observe that if $u = \varepsilon$ then $R = A$ and $K = \{\varepsilon\}$. Let $L_R = L \cap (R \times A^*)$ and $L_K = L \cap (K \times A^*)$. We first prove:

Lemma 18. $L = (L_R)^*L_K$.

Proof. It is easily seen that $L_RL = \{(w_1w_2, w_1'w_2') \mid (w_1, w_1') \in L_R, (w_2, w_2') \in L\}$ is included in L. It follows $L = (L_R)^*L$ that implies $(L_R)^*L_K \subseteq L$. For the converse inclusion, let $(w, w') \in L$; we consider two cases:

(i) $uw \notin A^+uA^+$. In this case, $w \in K \cup R$ that implies $(w, w') \in L_R \cup L_K \subseteq (L_R)^*L_K$.

(ii) $uw \in A^+uA^+$. In this case, $w = w_1w_2$ with $w_1 \in R$. So $uw_1 = \alpha u$ for some word α and by Corollary 2 it follows $uw' = w_1''w_2''$ for some words

w_1'', w_2'' with $uw_1 = \alpha u \xrightarrow{*} w_1''$ and $uw_2 \xrightarrow{*} uw_2''$. Observe that $w_1'' = uw_1'''$ for some word w_1''' since $u \to uf$, that implies $w' = w_1'''w_2''$. Now, from the inductive hypothesis over the length of w, we get $(w_2, w_2'') \in (L_R)^* L_K$. Moreover $(w_1, w_1''') \in L_R$ so $(w, w') = (w_1, w_1''')(w_2, w_2'') \in (L_R)^* L_K$.

□

For all word $e \in E$, we denote $R_e = u^{-1}S^*(\varphi(ue)))$ where φ is the function defined in Definition 1 and we define $U = \bigcup_{e \in E}(\{e\} \times R_e)$ and $V = (\{\varepsilon\} \times f^*)U^*\mathsf{Id}$. From Lemma 17, $R_e \in \mathsf{RAT}$ and, since E is finite, U and V are rational subsets of $A^* \times A^*$. We get:

Lemma 19. $L_K = V \cap (K \times A^*)$.

Proof. Since V is clearly included in L, we have the inclusion $V \cap (K \times A^*) \subseteq L_K$. Conversely, let $(w, w') \in L_K$. The proof is an induction over $|w|$, the length of the word w. If $w = \varepsilon$ then $w' \in f^*$, so $(w, w') \in \{\varepsilon\} \times f^* \subseteq V \cap (K \times A^*)$; if $|w| > 0$, since $w \in K$, it follows $uw \notin A^+uA^*$, so $u \neq \varepsilon$ and we consider two cases:

(i) uw' is not right-complete with respect to uw. In this case, $uw = uw_1y$ and $uw' = uw_1'y$ for some $y \neq \varepsilon$ and $uw_1 \xrightarrow{*} uw_1'$. From the inductives hypothesis, we get $(w_1, w_1') \in V$ that implies $(w, w') = (w_1, w_1')(y, y) \in V$.

(ii) uw' is right-complete with respect to uw. In this case, $uw \xrightarrow{*} zu \xrightarrow{*} uw'$ for some word z. Since $w \in K$ and $R_S = A^*uE^*$ from Proposition 8, it follows $w \in E^+$. That implies $w = ey$ with $e \in E$, $y \in E^*$, so $y \in K$ and we have from Lemma 11: $uey \xrightarrow{*} e'uy \xrightarrow{*} zu \xrightarrow{*} uw'$ with $e'u = \varphi(ue)$. From Proposition 10, it follows that $w' = w_1'w_2'$ with $ue \xrightarrow{*} e'u \xrightarrow{*} uw_1'$ and $uy \xrightarrow{*} uw_2'$. So, $(y, w_2') \in L_K$ that implies $(y, w_2') \in V$ from the inductive hypothesis and $(e, w_1') \in U$. Finally, we get $(w, w') \in UV \subseteq V$.

□

Lemma 20. $L_R = V(\{\varepsilon\} \times f^*) \cap (R \times A^*)$.

Proof. Let us first consider the case $u = \varepsilon$. In this case, $R = A$, $K = \{\varepsilon\}$, $L_R = \bigcup_{a \in A}(\{a\} \times f^*af^*)$, $U = \emptyset$ and $V = (\{\varepsilon\} \times f^*)\mathsf{Id}$. It follows that $L_R \subseteq V(\{\varepsilon\} \times f^*)$ so we have $L_R = V(\{\varepsilon\} \times f^*) \cap (R \times A^*)$. We now assume $u \neq \varepsilon$, so $u = u_0c$ for some word u_0 and some letter c. The inclusion $V(\{\varepsilon\} \times f^*) \cap (R \times A^*) \subseteq L_R$ is clear. To prove the converse inclusion, let $(w, w') \in L_R$. Since $w \in R$, $w = w_0c$ for some word w_0 and, since $uw_0c \xrightarrow{*} uw'$, w' can be factorized as $w' = w_1'cw_2'$ with $uw_0 \xrightarrow{*} uw_1'$ and $u \xrightarrow{*} uw_2'$ from Proposition 10. That implies $w_2' = f^i$ for some integer i so $w' = w_1'cf^i$. Now, since $w = w_0c \in R$, it follows $uw_0c \notin A^+uA^+$ that implies $uw_0 \notin A^+uA^*$ so $w_0 \in K$. As a consequence, we get $(w_0, w_1') \in L_K = V \cap (K \times A^*)$ so $(w_0c, w') = (w_0, w_1')(c, c)(\varepsilon, f^i)$ that is in $V(\{\varepsilon\} \times f^*)$. □

As a consequence of Lemma 18, Lemma 19 and Lemma 20, we get that L is a rational subset of $A^* \times A^*$ that implies that S is rational, so using also Proposition 14, we can state:

Proposition 15. *Let $S = \{u \mapsto uf\}$ be a prefixal one-rule rewrite system. The three following statements are equivalent:*

1. *S is rational*
2. *$S^*(u)$ is a rational language*
3. *$S^*(u) = uf^*$*

As a corollary, we get:

Corollary 3. *Let $S = \{u \mapsto uf\}$ be a prefixal one-rule string rewrite system. If $f \in a^+$ for some letter a then S is rational.*

Proof. Assume $f = a^i$ for some $i > 0$. Let us consider two cases:

(i) $ua^* \cap A^+uA^* = \emptyset$. In this case, $S^*(u) = uf^*$.
(ii) $ua^* \cap A^+uA^* \neq \emptyset$. That implies $u \in a^*$, so $S^*(u) = u(a^i)^* = uf^*$.

□

An important point is that Proposition 15 gives a characterization of rational prefixal one-rule rewrite systems that is decidable: indeed, given a prefixal system $S = \{u \mapsto uf\}$ one can decide whether $S^*(u)$ is a rational language (or equivalently whether $S^*(u) = uf^*$), thanks to the more general following result:

Proposition 16. *Let $S = \{u \mapsto v\}$ be a one-rule string rewrite system with $u \in \mathsf{F}(v)$. Then S is rational if and only if $S^*(u)$ is rational, moreover this property is decidable.*

Proof. Let $S = \{u \mapsto v\}$ with $u \in \mathsf{F}(v)$. There exist four cases:

(i) $u \notin \mathsf{LF}(v)$ and $u \notin \mathsf{RF}(v)$: in this case $v \in A^+uA^+$ so $S^*(u) \cap A^+uA^* \neq \emptyset$. From Lemma 16 it follows that $S^*(u)$ is not a rational language so S is not rational.
(ii) $v = uf = gu$ for some words f and g: from Lemma 15, we have to check the third statement of this lemma, that is clearly decidable.
(iii) $v = uf$ and $u \notin \mathsf{RF}(uf)$: the equivalence comes from Proposition 15. Moreover the rationality of $S^*(u)$ is equivalent to the emptiness of $S^*(u) \cap A^+uA^*$ from Lemma 16. To check whether $S^*(u) \cap A^+uA^*$ is empty or not is equivalent to check whether $uf^* \cap A^+uA^*$ is empty or not which is decidable.
(iv) $v = gu$ and $u \notin \mathsf{LF}(gu)$: this case is symmetric to the previous one.

□

7. Conclusion and Perspectives

We have proved in Proposition 11 that the image of a finite language by a prefixal one-rule rewrite system is always context-free, a property that is not satisfied by an arbitrary one-rule rewrite system as shown in Proposition 13.

This question can be extended to rational languages rather than finite ones: are prefixal one-rule rewrite systems $\mathsf{RAT/CF}$? And, in the case of a negative answer, is it possible to decide, given a prefixal one-rule rewrite system $S = \{u \mapsto uf\}$ whether it is $\mathsf{RAT/CF}$? More generally, we do not know whether prefixal one-rule rewrite systems preserve context-free languages like inverse match-bound rewrite systems.

Of course, some of these questions deserve to be studied in the more general case of one-rule rewrite systems: is it possible to decide, given a one-rule rewrite system $S = \{u \mapsto v\}$

$\mathcal{Q}_1$) whether, S^* preserves rational languages ?

$\mathcal{Q}_2$) whether for all word w, $S^*(w)$ is a rational language ?

$\mathcal{Q}_3$) whether, given a word w, $S^*(w)$ is a rational language ?

$\mathcal{Q}_4$) whether, S^* preserves context-free languages ?

$\mathcal{Q}_5$) whether for all word w, $S^*(w)$ is a context-free language ?

$\mathcal{Q}_6$) whether, given a word w, $S^*(w)$ is a context-free language ?

Note that the following one-rule rewrite system $S = \{ba \mapsto a^2b\}$ gives an interesting example of a system where for all word w, $S^*(w)$ is a context-free language, but where $S^*(b^*a) \cap a^*b^* = \{a^{2^p}b^p \mid p \geq 0\}$ is not a context-free language.

We also remark that, in the case of prefixal one-rule rewrite systems, only $\mathcal{Q}_4$ is an open question. Indeed, we have positively answered to $\mathcal{Q}_1$ and $\mathcal{Q}_2$ and, using the technics of Section 6, it is possible to positively answer to $\mathcal{Q}_3$ by proving that $S^*(w)$ is a rational language if and only if $S^*(u)$ is rational or $w \notin A^*uA^*$. Trivial positive answer to $\mathcal{Q}_5$ and $\mathcal{Q}_6$ is a direct consequence of Proposition 11.

Proposition 16 gives a decidable characterization of rational prefixal one-rule rewrite systems. We observe that we are unable to give such a decidable characterization for one-rule length-preserving rational rewrite systems $S = \{u \mapsto v\}$, that is in the case when $|u| = |v|$ and $u \neq v$: indeed, it is conjectured that a one-rule length-preserving rewrite system is a rational transduction if and only if the left-hand side u and the right-hand side v of the rule of the system are not quasi-conjugate or are equal, that means if u and v are distinct, there do not exist words x, y and z such that $u = xyz$ and $v = zyx$ ([TS00, LR14, LR15]). This conjecture is still open. More generally, one can wonder whether this property is decidable for an arbitrary one-rule rewrite system.

Acknowledgments

The authors thank the anonymous reviewer for his very careful reading and his many insightful comments and suggestions.

References

[Ber79] Jean Berstel. *Transductions and context-free languages*, volume 38 of *Teubner Studienbücher : Informatik*. Teubner, 1979.

[BO93] Ronald V. Book and Friedrich Otto. *String-Rewriting Systems*. Texts and Monographs in Computer Science. Springer, 1993.

[CR90] Mireille Clerbout and Yves Roos. Semi-commutations and algebraic languages. In Christian Choffrut and Thomas Lengauer, editors, *STACS 90*, volume 415 of *Lecture Notes in Computer Science*, pages 82–94. Springer Berlin / Heidelberg, 1990.

[Der05] Nachum Dershowitz. Open. closed. open. In Jürgen Giesl, editor, *RTA*, volume 3467 of *Lecture Notes in Computer Science*, pages 376–393. Springer, 2005.

[Ges02] Alfons Geser. Decidability of termination of grid string rewriting rules. *SIAM J. Comput.*, 31(4):1156–1168, 2002.

[Ges03] Alfons Geser. Termination of string rewriting rules that have one pair of overlaps. In Robert Nieuwenhuis, editor, *RTA*, volume 2706 of *Lecture Notes in Computer Science*, pages 410–423. Springer, 2003.

[GHW04] Alfons Geser, Dieter Hofbauer, and Johannes Waldmann. Match-bounded string rewriting systems. *Appl. Algebra Eng. Commun. Comput.*, 15(3-4):149–171, 2004.

[GHW05] Alfons Geser, Dieter Hofbauer, and Johannes Waldmann. Termination proofs for string rewriting systems via inverse match-bounds. *J. Autom. Reasoning*, 34(4):365–385, 2005.

[GZ99] Alfons Geser and Hans Zantema. Non-looping string rewriting. *ITA*, 33(3):279–302, 1999.

[HW04] Dieter Hofbauer and Johannes Waldmann. Deleting string rewriting systems preserve regularity. *Theor. Comput. Sci.*, 327(3):301–317, 2004.

[Jan85] Matthias Jantzen. A note on a special one-rule semi-thue system. *Inf. Process. Lett.*, 21(3):135–140, 1985.

[Jan88] M. Jantzen. *Confluent String Rewriting*. EATCS monographs on theoretical computer science. Springer-Verlag, 1988.

[Kur90] Winfried Kurth. *Termination und Konfluenz von Semi–Thue–Systemen mit nur einer Regel*. PhD thesis, Technische Universität Clausthal, 1990.

[Leu08] Peter Leupold. On regularity-preservation by string-rewriting systems. In Carlos Martín-Vide, Friedrich Otto, and Henning Fernau, editors, *Language and Automata Theory and Applications, Second International Conference, LATA 2008, Tarragona, Spain, March 13-19, 2008. Revised Papers*, volume 5196 of *Lecture Notes in Computer Science*, pages 345–356. Springer, 2008.

[LR12] Michel Latteux and Yves Roos. On one-rule grid semi-thue systems. *Fundam. Inform.*, 116(1-4):189–204, 2012.

[LR14] Michel Latteux and Yves Roos. One-rule length-preserving rewrite systems and rational transductions. *RAIRO - Theor. Inf. and Applic.*, 48(2):149–171, 2014.

[LR15] Michel Latteux and Yves Roos. A canonical automaton for one-rule length-preserving string rewrite systems. *Inf. Comput.*, 244:203–228, 2015.

[McN01] Robert McNaughton. Semi-thue systems with an inhibitor. *J. Autom. Reasoning*, 26(4):409–431, 2001.

[Sén90] Géraud Sénizergues. Some decision problems about controlled rewriting systems. *Theor. Comput. Sci.*, 71(3):281–346, 1990.

[Sén95] Géraud Sénizergues. Formal languages & word-rewriting. In Hubert Comon and Jean-Pierre Jounnaud, editors, *Term Rewriting: French Spring School of Theoretical Computer Science Font Romeux, France, May 17–21, 1993 Advanced Course*, pages 75–94, Berlin, Heidelberg, 1995. Springer Berlin Heidelberg.

[TS00] Alain Terlutte and David Simplot. Iteration of rational transductions. *ITA*, 34(2):99–130, 2000.

[Wra90] Celia Wrathall. Confluence of one-rule thue systems. In Klaus U. Schulz, editor, *Word Equations and Related Topics, First International Workshop, IWWERT '90, Tübingen, Germany, October 1-3, 1990, Proceedings*, volume 572 of *Lecture Notes in Computer Science*, pages 237–246. Springer, 1990.