

Undecidable Properties on Length-Two String Rewriting Systems

Masahiko Sakai

Graduate School of Information Science, Nagoya University, Yi Wang
Graduate School of Arts and Science, The University of Tokyo

Abstract

Length-two string rewriting systems are length preserving string rewriting systems that consist of length-two rules. This paper shows that confluence, termination, left-most termination and right-most termination are undecidable properties for length-two string rewriting systems. This results mean that these properties are undecidable for the class of linear term rewriting systems in which depth-two variables are allowed in both-hand sides of rules.

1 Introduction

Confluence and termination are both generally undecidable for term rewriting systems (TRSs) and for string rewriting systems. Hence several decidable classes have been studied. Confluence is a decidable property for terminating TRSs [11], ground TRSs [14], linear shallow TRSs [6] and shallow right-linear TRSs [7]. Termination is a decidable property for ground TRSs [9], right ground TRSs [3], TRSs that consist of right-ground rules, collapsing rules and shallow right-linear rules [8], and related class of shallow left-linear TRSs [16].

There are also results on undecidable classes. Confluence is an undecidable property for semi-constructor TRSs [12] and flat TRSs [10, 13]. Caron showed that termination is an undecidable property for length preserving string rewriting systems [2].

String rewriting systems (SRSs) are said to be length preserving if the left-hand side and the right-hand side of each rule have the same length. Especially they are length-two systems if all of the lengths are two. In this paper, we show confluence, termination, left-most termination and right-most termination are undecidable properties for length-two SRSs. Firstly we show those for length preserving SRSs by reducing the Post's correspondence problem, which is known to be undecidable, to termination problem and to confluence problem for length preserving SRSs. Then we give a transformation of length preserving SRSs to length-two SRSs that preserves both confluence and termination properties.

The class of length-two SRSs is a subclass of linear term rewriting systems in which depth-two variables are allowed in both-hand sides of rules. Thus the undecidability for these class of term rewriting systems also obtained. In that sence, the undecidability results in this paper shed new light on the borderline between decidability and undecidability for TRSs.

2 Preliminaries

Let Σ be an alphabet. A *string rewrite rule* is a pair of strings $l, r \in \Sigma^*$, denoted by $l \rightarrow r$. A finite set of string rewrite rules is called a *string rewriting system* (SRS). A string is called a *redex* if it is the left-hand side of a rule. An SRS \mathcal{R} induces a *rewrite step* relation $\xrightarrow{\mathcal{R}}$ defined as $s \xrightarrow{\mathcal{R}} t$ if there exist $u, v \in \Sigma^*$ and a rule $l \rightarrow r$ in \mathcal{R} such that $s = ulv$ and $t = urv$. Especially it is *left-most* (resp. *right-most*) if l is the left-most (resp. right-most) redex in s . We use $\xrightarrow[\mathcal{R}]{}^+$ for the transitive closure of $\xrightarrow{\mathcal{R}}$ and $\xrightarrow[\mathcal{R}]{}^*$ for the reflexive-transitive closure of $\xrightarrow{\mathcal{R}}$. We use $\xleftrightarrow[\mathcal{R}]{}^k$ for $\xleftarrow[\mathcal{R}]{} \cup \xrightarrow[\mathcal{R}]{}^k$. We write $\xrightarrow[\mathcal{R}]{}^k$ for the relation with k rewrite steps. A (possibly infinite) sequence $s_0 \xrightarrow[\mathcal{R}]{} s_1 \xrightarrow[\mathcal{R}]{} \cdots$ is called a reduction sequence.

We say that string s is *terminating* if every reduction sequence starting from s is finite. We say that strings s_1 and s_2 are *joinable* if $s_1 \xrightarrow[\mathcal{R}]{}^* s \xleftarrow[\mathcal{R}]{}^* s_2$ for some s , denoted by $s_1 \downarrow_{\mathcal{R}} s_2$. A string s is *confluent* if $s_1 \downarrow_{\mathcal{R}} s_2$ for any $s_1 \xleftarrow[\mathcal{R}]{}^* s \xrightarrow[\mathcal{R}]{}^* s_2$. An SRS \mathcal{R} is *confluent* (*terminating*) if all strings are confluent (terminating).

In this paper, the notation $|u|$ represents the length of string u . The notation $\underbrace{a \cdots a}_m$ denotes the string that consists of m symbols of a . We refer $\{r \rightarrow l \mid l \rightarrow r \in \mathcal{R}\}$ by \mathcal{R}^{-1} .

Now we recall Post's correspondence problem (PCP), which is known to be undecidable.

Definition 2.1 *An instance of PCP is a finite set $P \subseteq \mathcal{A}^* \times \mathcal{A}^*$ of finite pairs of non-empty strings over an alphabet \mathcal{A} with at least two symbols. A solution of P is a string w such that*

$$w = u_1 \cdots u_k = v_1 \cdots v_k$$

for some $(u_i, v_i) \in P$. The Post's correspondence problem (PCP) is a problem to decide whether such a solution exists or not.

Example 2.2 The set $P = \{(ab, a), (a, ba)\}$ is an instance of PCP over $\{a, b\}$. It has a solution $aba = u_1 u_2 = v_1 v_2$ with $(u_1, v_1) = (ab, a)$, $(u_2, v_2) = (a, ba)$.

Theorem 2.3 ([15]) *PCP is undecidable.*

3 Length preserving SRSs and undecidability of their termination

Definition 3.1 *An SRS \mathcal{R} is said to be length preserving if $|l| = |r|$ for every rule $l \rightarrow r$ in \mathcal{R} .*

Since there is a finite number of rules, the number of different symbols appearing in the rules is finite, and fixed for them. Hence the number of strings with a given length is also finite. Thus the decidability of the following problems for length preserving SRSs trivially follows.

1. *Reachability problem:* problem to decide $s \xrightarrow[\mathcal{R}]{}^* t$ for given strings s and t and a SRS \mathcal{R} .
2. *String-confluence problem:* problem to decide confluence of s for a given string s and a SRS \mathcal{R} .
3. *String-termination problem:* problem to decide termination of s for a given string s and a SRS \mathcal{R} .

In this section we argue about the undecidability of termination, right-most termination and left-most termination for length preserving SRSs. As stated in the introduction, Caron showed the undecidability in [2]. Moreover the proof works also for right-most termination and left-most termination because there is only one redex in each string that corresponds to a correct automata configuration. Nevertheless we give an alternative proof from the following reasons:

- Caron's proof composed of two stages; the first stage gives an algorithm that reduce PCP into the uniform halting problem for linear-bounded automata and the second stage gives an algorithm reducing the uniform halting problem into the termination problem for length preserving SRSs. On the other hand, we give a proof by reducing the Post's correspondence problem into termination problem of SRSs directly.
- The SRS \mathcal{T}_P given in this section is rather straightforward and easy to understand. This helps the understanding of SRS \mathcal{C}_P given in the next section, which is more difficult although it is just a variant of \mathcal{T}_P .

As a preparation for giving the transformation, we introduce a kind of null symbol $-$ and an equal length representation of each pair in instances of PCP. Let $P = \{(u_1, v_1), \dots, (u_n, v_n)\}$ be an instance of PCP over \mathcal{A} .

$$\begin{aligned} \bar{P} &= \{(u, \underbrace{v \dots -}_{|m|}) \mid (u, v) \in P \text{ and } |u| - |v| = m \geq 0\} \\ &\cup \{(u, \underbrace{- \dots v}_{|m|}) \mid (u, v) \in P \text{ and } |u| - |v| = m < 0\} \end{aligned}$$

We write $\bar{\mathcal{A}}$ for $\mathcal{A} \cup \{-\}$. We define an equivalence relation $\sim_{\subseteq} (\bar{\mathcal{A}})^* \times (\bar{\mathcal{A}})^*$ as identity relation with ignoring all null symbols $-$, that is $u \sim v$ if and only if

$\hat{u} = \hat{v}$ where \hat{u} and \hat{v} denote the strings obtained from u and v by removing all $-$ s respectively.

Example 3.2 For an instance $P = \{(ab, a), (a, ba)\}$ of PCP, we have $\bar{P} = \{(ab, a-), (a-, ba)\}$. The solution corresponds to $u_1 u_2 = ab a- \sim a- ba = v_1 v_2$ for $(u_1, v_1), (u_2, v_2) \in \bar{P}$.

We use symbols like $0_{a'}^b$, where 0 is called the *tag* of the symbol and a is called the *first subscript* of the symbol, b the *second*, a' the *third* and b' the *fourth*. We code the solution of the previous example into $\tilde{0}_{a-}^a \tilde{0}_{b-}^b \tilde{0}_{a-}^a \tilde{0}_{b-}^a$.

For an easy handling of strings that consist of such symbols, we introduce a notation defined as

$$(X_1 \cdots X_k)_{\substack{b_1 \cdots b_k \\ a'_1 \cdots a'_k \\ b'_1 \cdots b'_k}}^{a_1 \cdots a_k} = X_1_{a'_1}^{b_1} \cdots X_k_{a'_k}^{b_k}.$$

For example the above solution is denoted by $(\tilde{0}\tilde{0})_{a-}^{ab} (\tilde{0}\tilde{0})_{b-}^{a-}$ or $(\tilde{0}\tilde{0}\tilde{0}\tilde{0})_{aba-}^{aba-}$. Note that the lengths of the strings in those subscripts are the same whenever we use this notation.

The first and the second subscripts keep a candidate of solutions of P in equal length representation and will never be changed by reductions. The third and the fourth subscripts are used as working area for checking whether the candidate is a solution or not.

We relate a solution of the given instance of PCP with a loop in an infinite reduction sequence:

$$\begin{aligned} & \Xi_0 (\hat{0}\hat{0} \cdots \hat{0})_{v_1}^{u_1} \cdots (\hat{0}\hat{0} \cdots \hat{0})_{v_k}^{u_k} \Psi_0 \xrightarrow[T_P]{*} \Xi_2 (\hat{2}\hat{2} \cdots \hat{2})_{w_1}^{u_1} \cdots (\hat{2}\hat{2} \cdots \hat{2})_{w_k}^{u_k} \Psi_2 \\ & \xrightarrow[T_P]{*} \Xi_0 (\hat{0}\hat{0} \cdots \hat{0})_{v_1}^{u_1} \cdots (\hat{0}\hat{0} \cdots \hat{0})_{v_k}^{u_k} \Psi_0. \end{aligned}$$

1. The former part checks whether $u_1 \cdots u_k \sim v_1 \cdots v_k$ by using the third and the fourth subscripts as working area.
2. The latter part checks whether $(u_1, v_1), \dots, (u_n, v_n) \in \bar{P}$ and initializes the working area.

Definition 3.3 Let P be an instance of PCP over \mathcal{A} . The SRS \mathcal{T}_P over Σ obtained from P is defined as follows, where individual rules are shown in Figure 1.

$$\begin{aligned} \Sigma &= \{\Xi_i, \Psi_i \mid i \in \{0, 1, 2\}\} \cup \Sigma_c \\ \Sigma_c &= \left\{ n_{x_3}^{x_1}, \hat{n}_{x_3}^{x_2}, n_{x_3}^{x_2}, \hat{n}_{x_3}^{x_1} \mid x_i \in \bar{\mathcal{A}}, n \in \{0, 1, 2\} \right\} \\ \mathcal{T}_P &= \alpha_1 \cup \beta_1 \cup \gamma_1 \cup \alpha_2 \cup \beta_2 \cup \gamma_2 \cup \delta_2 \end{aligned}$$

$$\begin{aligned}
\alpha_1 &= \left\{ (\tilde{1}2 \cdots 2)_{u', v'}^u \Psi_2 \rightarrow (\tilde{0}0 \cdots 0)_{u', v'}^u \Psi_0 \mid (u, v) \in \bar{P}, u', v' \in (\bar{\mathcal{A}})^* \right\} \\
\beta_1 &= \left\{ (\tilde{1}2 \cdots 2)_{u', v'}^u \tilde{2}_{x_3, x_4}^{x_1} \rightarrow (\tilde{0}0 \cdots 0)_{u', v'}^u (\tilde{1})_{x_3, x_4}^{x_1} \mid (u, v) \in \bar{P}, u', v' \in (\bar{\mathcal{A}})^*, x_j \in \bar{\mathcal{A}} \right\} \\
\gamma_1 &= \left\{ \Xi_2 \tilde{2}_{x_3, x_4}^{x_1} \rightarrow \Xi_0 \tilde{1}_{x_3, x_4}^{x_1} \mid x_j \in \bar{\mathcal{A}} \right\} \\
\alpha_2 &= \left\{ 0_{x_3}^{x_1} \Psi_0 \rightarrow \underline{2}_{x_3}^{x_1} \Psi_2 \mid x_j \in \bar{\mathcal{A}} \right\} \\
\beta_2 &= \left\{ 0_{x_3}^{x_1} \underline{2}_{y_3}^{y_1} \rightarrow \underline{2}_{x_3}^{x_1} \underline{2}_{y_3}^{y_1}, \tilde{0}_{x_3}^{x_1} \underline{2}_{y_3}^{y_1} \rightarrow \underline{2}_{x_3}^{x_1} \underline{2}_{y_3}^{y_1}, 0_{x_3}^{x_1} \underline{2}_{y_3}^{y_1} \rightarrow \underline{2}_{x_3}^{x_1} \underline{2}_{y_3}^{y_1} \mid x_j, y_j \in \bar{\mathcal{A}} \right\} \\
\gamma_2 &= \left\{ \Xi_0 \underline{2}_{x_3}^{x_1} \rightarrow \Xi_2 \tilde{2}_{x_3}^{x_1} \mid x_j \in \bar{\mathcal{A}} \right\} \\
\delta_2 &= \left\{ X_{x_4}^{x_1} Y_{y_4}^{y_1} \rightarrow X_{x_4}^{x_1} Y_{y_4}^{y_1}, X_{x_3}^{x_1} Y_{y_3}^{y_1} \rightarrow X_{x_3}^{x_1} Y_{y_3}^{y_1} \mid x_j, y_j \in \bar{\mathcal{A}}, z \in \mathcal{A}, X, Y \in \{0, \tilde{0}\} \right\}
\end{aligned}$$

Figure 1: Rules in \mathcal{T}_P

Example 3.4 Consider the instance $P = \{(a, ba), (ab, a)\}$ of PCP. Rules α_1, β_1 depend on P and the other rules depend only on the alphabet \mathcal{A} .

$$\begin{aligned}
\alpha_1 &= \left\{ (\tilde{1}2)_{x_2 y_2}^{b a} \Psi_2 \rightarrow (\tilde{0}0)_{a-}^{a-} \Psi_0, (\tilde{1}2)_{x_2 y_2}^{a b} \Psi_2 \rightarrow (\tilde{0}0)_{a-}^{a-} \Psi_0 \mid x_i, y_i \in \bar{\mathcal{A}} \right\} \\
\beta_1 &= \left\{ (\tilde{1}2)_{x_2 y_2}^{a-} \tilde{2}_{z_3, z_4}^{z_1} \rightarrow (\tilde{0}0)_{a-}^{a-} \tilde{1}_{z_3, z_4}^{z_1}, (\tilde{1}2)_{x_2 y_2}^{a b} \tilde{2}_{z_3, z_4}^{z_1} \rightarrow (\tilde{0}0)_{a-}^{a-} \tilde{1}_{z_3, z_4}^{z_1} \mid x_i, y_i, z_i \in \bar{\mathcal{A}} \right\}
\end{aligned}$$

\mathcal{T}_P is not terminating since we can construct an infinite reduction sequence. We start with a string $\Xi_0(\tilde{0}0)_{a-}^{a-}(\tilde{0}0)_{a-}^{a-}\Psi_0$. Rules in δ_2 move null symbols in the third or the fourth subscripts into the tail:

$$\Xi_0(\tilde{0}0)_{a-}^{a-}(\tilde{0}0)_{a-}^{a-}\Psi_0 \xrightarrow{\delta_2} \Xi_0(\tilde{0}0)_{a-}^{a-}(\tilde{0}0)_{a-}^{a-}\Psi_0 \xrightarrow{\delta_2} \Xi_0(\tilde{0}0)_{a-}^{a-}(\tilde{0}0)_{a-}^{a-}\Psi_0.$$

Rules in $\alpha_2 \cup \beta_2 \cup \gamma_2$ check in right-to-left order that the third and the fourth subscripts are the same:

$$\begin{aligned}
&\Xi_0(\tilde{0}0)_{a-}^{a-}(\tilde{0}0)_{a-}^{a-}\Psi_0 \xrightarrow{\alpha_2} \Xi_0(\tilde{0}0)_{a-}^{a-}(\tilde{0}\underline{2})_{a-}^{a-}\Psi_2 \xrightarrow{\beta_2} \Xi_0(\tilde{0}0)_{a-}^{a-}(\tilde{2}\underline{2})_{a-}^{a-}\Psi_2 \\
&\xrightarrow{\beta_2} \Xi_0(\tilde{0}\underline{2})_{a-}^{a-}(\tilde{2}\underline{2})_{a-}^{a-}\Psi_2 \xrightarrow{\beta_2} \Xi_0(\tilde{2}\underline{2})_{a-}^{a-}(\tilde{2}\underline{2})_{a-}^{a-}\Psi_2 \xrightarrow{\gamma_2} \Xi_2(\tilde{2}\underline{2})_{a-}^{a-}(\tilde{2}\underline{2})_{a-}^{a-}\Psi_2.
\end{aligned}$$

Rules in $\gamma_1 \cup \beta_1 \cup \alpha_1$ check in left-to-right order that the first and the second subscripts consist of pairs in \bar{P} and copy the first subscript to the third and the second to the fourth respectively:

$$\begin{aligned}
&\Xi_2(\tilde{2}\underline{2})_{a-}^{a-}(\tilde{2}\underline{2})_{a-}^{a-}\Psi_2 \xrightarrow{\gamma_1} \Xi_0(\tilde{1}2)_{a-}^{a-}(\tilde{2}\underline{2})_{a-}^{a-}\Psi_2 \xrightarrow{\beta_1} \Xi_0(\tilde{0}0)_{a-}^{a-}(\tilde{1}2)_{a-}^{a-}\Psi_2 \\
&\xrightarrow{\alpha_1} \Xi_0(\tilde{0}0)_{a-}^{a-}(\tilde{0}0)_{a-}^{a-}\Psi_0.
\end{aligned}$$

Obviously \mathcal{T}_P is length preserving. The proof of the following lemma is found in Section 5.

Lemma 3.5 *Let P be an instance of PCP. Then the following properties are equivalent:*

1. P has a solution.
2. \mathcal{T}_P is not right-most terminating.
3. \mathcal{T}_P is not left-most terminating.
4. \mathcal{T}_P is not terminating.

Theorem 3.6 *Termination, right-most termination and left-most termination are undecidable properties for length preserving SRSs.*

Proof. We assume that termination (right-most termination, left-most termination) of length preserving SRSs is decidable. Then it follows from Lemma 3.5 that PCP is decidable, which contradicts to Theorem 2.3. \square

4 Undecidability of confluence for length preserving SRSs

We modify the construction of SRS in the last section. In contrast to the SRS \mathcal{T}_P , which works sequentially, the SRS \mathcal{C}_P works in parallel, that is, a solution of a given instance of PCP is related with the following two reduction sequences

$$\begin{aligned} & \Xi_0(\tilde{00}\cdots 0)_{\substack{u_1 \\ v_1}}^{u_1} \cdots (\tilde{00}\cdots 0)_{\substack{u_k \\ v_k}}^{u_k} \Psi_0 \xrightarrow[\mathcal{C}_P]{*} \Xi_2(\tilde{22}\cdots 2)_{\substack{u_1 \\ w_1}}^{u_1} \cdots (\tilde{22}\cdots 2)_{\substack{u_k \\ w_k}}^{u_k} \Psi_2 \\ & \Xi_0(\tilde{00}\cdots 0)_{\substack{u_1 \\ v_1}}^{u_1} \cdots (\tilde{00}\cdots 0)_{\substack{u_k \\ v_k}}^{u_k} \Psi_0 \xrightarrow[\mathcal{C}_P]{*} \Xi_1(\tilde{11}\cdots 1)_{\substack{u_1 \\ v_1}}^{u_1} \cdots (\tilde{11}\cdots 1)_{\substack{u_k \\ v_k}}^{u_k} \Psi_1 \end{aligned}$$

that demonstrate its non-confluence.

1. The former reduction checks whether $u_1 \cdots u_k \sim v_1 \cdots v_k$ by using the third and the fourth subscripts as working area.
2. The latter reduction checks whether $(u_1, v_1), \dots, (u_n, v_n) \in \bar{P}$ and checks the working area is correctly initialized.

In case of P has no solution, \mathcal{C}_P must be confluent, which makes the design of \mathcal{C}_P difficult.

Definition 4.1 *Let P be an instance of PCP over \mathcal{A} . The SRS \mathcal{C}_P over Σ obtained from P is defined as follows:*

$$\begin{aligned} \mathcal{C}_P &= \Theta \cup \Phi, \\ \Theta &= \Theta_1 \cup \Theta_2, \\ \Theta_1 &= \alpha'_1 \cup \beta'_1 \cup (\alpha'_1 \cup \beta'_1)^{-1}, \\ \Theta_2 &= \alpha_2 \cup \beta_2 \cup \delta_2 \cup \epsilon_2 \cup (\alpha_2 \cup \beta_2 \cup \delta_2 \cup \epsilon_2)^{-1}, \\ \Phi &= \gamma'_1 \cup \gamma_2 \end{aligned}$$

where rules α_2 , β_2 , δ_2 and γ_2 are shown in Figure 1 and the other rules are shown in Figure 2.

$$\begin{aligned}
\alpha'_1 &= \left\{ (\tilde{0}0 \cdots 0)_{\substack{u \\ v}}^u \Psi_0 \rightarrow (\tilde{1}1 \cdots 1)_{\substack{u \\ v}}^u \Psi_1 \mid (u, v) \in \bar{P} \right\} \\
\beta'_1 &= \left\{ (\tilde{0}0 \cdots 0)_{\substack{u \\ v}}^u \tilde{1}_{\substack{x_1 \\ x_2}}^{x_1} \rightarrow (\tilde{1}1 \cdots 1)_{\substack{u \\ v}}^u \tilde{1}_{\substack{x_1 \\ x_2}}^{x_1} \mid (u, v) \in \bar{P}, x_i \in \bar{\mathcal{A}} \right\} \\
\gamma'_1 &= \left\{ \Xi_0 \tilde{1}_{\substack{x_1 \\ x_2}}^{x_1} \rightarrow \Xi_1 \tilde{1}_{\substack{x_1 \\ x_2}}^{x_1} \mid x_i \in \bar{\mathcal{A}} \right\} \\
\epsilon_2 &= \left\{ X_{\substack{x_1 \\ x_2}}^{x_1} Y_{\substack{y_1 \\ y_2}}^{y_1} \rightarrow X_{\substack{x_1 \\ x_2}}^{x_1} Y_{\substack{y_1 \\ y_2}}^{y_1}, X_{\substack{x_1 \\ x_2}}^{x_1} Y_{\substack{y_1 \\ y_2}}^{y_1} \rightarrow X_{\substack{x_1 \\ x_2}}^{x_1} Y_{\substack{y_1 \\ y_2}}^{y_1} \mid x_j, y_j \in \bar{\mathcal{A}}, z \in \mathcal{A}, X, Y \in \{2, \tilde{2}\} \right\}
\end{aligned}$$

Figure 2: Rules in \mathcal{C}_P

Remark that the reductions by Θ -rules are symmetric, that is to say, $s \xrightarrow{\Theta} t$ if and only if $t \xrightarrow{\Theta} s$, which plays an important role to make \mathcal{C}_P confluent if P has no solution.

Example 4.2 Let $P = \{(a, ba), (ab, a)\}$ be an instance of PCP. Rules α'_1, β'_1 depends on P and the other rules depend only on the alphabet \mathcal{A} .

$$\begin{aligned}
\alpha'_1 &= \left\{ (\tilde{0}0)_{\substack{ba \\ ba}}^{ab} \Psi_0 \rightarrow (\tilde{1}1)_{\substack{ba \\ ba}}^{ab} \Psi_1, (\tilde{0}0)_{\substack{ab \\ ab}}^{ab} \Psi_0 \rightarrow (\tilde{1}1)_{\substack{ab \\ ab}}^{ab} \Psi_1 \right\} \\
\beta'_1 &= \left\{ (\tilde{0}0)_{\substack{ba \\ ba}}^{ab} \tilde{1}_{\substack{x_1 \\ x_2}}^{x_1} \rightarrow (\tilde{1}1)_{\substack{ba \\ ba}}^{ab} \tilde{1}_{\substack{x_1 \\ x_2}}^{x_1}, (\tilde{0}0)_{\substack{ab \\ ab}}^{ab} \tilde{1}_{\substack{x_1 \\ x_2}}^{x_1} \rightarrow (\tilde{1}1)_{\substack{ab \\ ab}}^{ab} \tilde{1}_{\substack{x_1 \\ x_2}}^{x_1} \mid x_i \in \bar{\mathcal{A}} \right\}
\end{aligned}$$

We can show that \mathcal{C}_P is not confluent since we have non-joinable branches.

$$\begin{aligned}
&\Xi_0 (\tilde{0}0)_{\substack{ab \\ ab}}^{ab} (\tilde{0}0)_{\substack{ba \\ ba}}^{ab} \Psi_0 \xrightarrow{\alpha'_1} \Xi_0 (\tilde{0}0)_{\substack{ab \\ ab}}^{ab} (\tilde{1}1)_{\substack{ba \\ ba}}^{ab} \Psi_1 \xrightarrow{\beta'_1} \Xi_0 (\tilde{1}1)_{\substack{ab \\ ab}}^{ab} (\tilde{1}1)_{\substack{ba \\ ba}}^{ab} \Psi_1 \\
&\xrightarrow{\gamma'_1} \Xi_1 (\tilde{1}1)_{\substack{ab \\ ab}}^{ab} (\tilde{1}1)_{\substack{ba \\ ba}}^{ab} \Psi_1, \\
&\Xi_0 (\tilde{0}0)_{\substack{ab \\ ab}}^{ab} (\tilde{0}0)_{\substack{ba \\ ba}}^{ab} \Psi_0 \xrightarrow{\delta_2 \cup \alpha_2 \cup \beta_2 \cup \gamma_2} \Xi_2 (\tilde{2}2)_{\substack{ab \\ ab}}^{ab} (\tilde{2}2)_{\substack{ba \\ ba}}^{ab} \Psi_2.
\end{aligned}$$

Note that the detail of the latter sequence is found in Example 3.4.

Obviously \mathcal{C}_P is length preserving. The proof of the following main lemma is found in the next section.

Lemma 4.3 *Let P be an instance of PCP. Then, P has a solution if and only if \mathcal{C}_P is not confluent.*

Theorem 4.4 *Confluence of length preserving SRSs is an undecidable property.*

Proof. We assume that the problem is decidable. Then it follows from Lemma 4.3 that PCP is decidable, which contradicts to Theorem 2.3. \square

5 Proofs

Every occurrence of the symbols Ξ_1 , Ξ_2 and Ξ_3 in rules are right-most positions in both-hand sides. Moreover, for every rule, Ξ_i appears in the left-hand side if and only if Ξ_j appears in the right-hand side. Hence we can separate any reduction sequence having a symbol Ξ_i into two reduction sequences by cutting each string at the Ξ_j occurrence. Symbols Ψ_i also have the similar property. Therefore the following proposition holds.

Proposition 5.1 *Let \mathcal{R} be \mathcal{T}_P or \mathcal{C}_P obtained from an instance P of PCP. For any $i \in \{0, 1, 2\}$ and $S_1, S_2, S \in \Sigma^*$, the followings hold:*

- (a) *If $S_1\Xi_iS_2 \xrightarrow{\mathcal{R}} S$, then $(S = S'_1\Xi_iS_2) \wedge (S_1 \xrightarrow{\mathcal{R}} S'_1)$ or $(S = S_1\Xi_jS'_2) \wedge (\Xi_iS_2 \xrightarrow{\mathcal{R}} \Xi_jS'_2)$ for some $S'_1, S'_2 \in \Sigma^*$ and $j \in \{0, 1, 2\}$.*
- (b) *If $S_1\Xi_iS_2 \xrightarrow{\mathcal{R}}^* S$, then $S = S'_1S'_2$, $S_1 \xrightarrow{\mathcal{R}}^* S'_1$ and $\Xi_iS_2 \xrightarrow{\mathcal{R}}^* S'_2$ for some $S'_1 \in \Sigma^*$ and non-empty $S'_2 \in \Sigma^*$.*
- (c) *If $S_1\Psi_iS_2 \xrightarrow{\mathcal{R}} S$, then $(S = S'_1\Psi_jS_2) \wedge (S_1\Psi_i \xrightarrow{\mathcal{R}} S'_1\Psi_j)$ or $(S = S_1\Psi_iS'_2) \wedge (S_2 \xrightarrow{\mathcal{R}} S'_2)$ for some $S'_1, S'_2 \in \Sigma^*$ and $j \in \{0, 1, 2\}$.*
- (d) *If $S_1\Psi_iS_2 \xrightarrow{\mathcal{R}}^* S$, then $S = S'_1S'_2$, $S_1\Psi_i \xrightarrow{\mathcal{R}}^* S'_1$ and $S_2 \xrightarrow{\mathcal{R}}^* S'_2$ for some $S'_2 \in \Sigma^*$ and non-empty $S'_1 \in \Sigma^*$.*

Proof. We prove (a). Let $S_1\Xi_iS_2 \xrightarrow{\mathcal{R}} S$. The only interesting case is that the redex in the rewrite step contains the displayed symbol Ξ_i . Then one of γ_1 -rules, γ_2 -rules or γ'_1 -rules is applied. From the construction of the rules, we have $S = S_1\Xi_jS'_2$ and $\Xi_iS_2 \xrightarrow{\mathcal{R}} \Xi_jS'_2$ for some $S'_2 \in \Sigma^*$ and $j \in \{0, 1, 2\}$.

The claim (b) is easily proved by induction on the number k of the rewrite steps in $S_1\Xi_iS_2 \xrightarrow{\mathcal{R}}^* S$. For (c) and (d), the proofs are similar to (a) and (b) respectively. \square

We say a string over Σ is *normal* if it is in one of the following three forms:

$$(p1) \Xi_i\chi, \quad (p2) \chi\Psi_j, \quad (p3) \Xi_i\chi\Psi_j,$$

where $\chi \in (\Sigma_c)^*$, $i, j \in \{0, 1, 2\}$.

We prepare a measure for the proof of the next lemma. For a non-empty string $X_1 \cdots X_n$ over Σ , we define $\|X_1 \cdots X_n\|$ by the summation of the number of occurrences of Ξ_i symbols in $X_2 \cdots X_n$ and the number of occurrences of Ψ_i symbols in $X_1 \cdots X_{n-1}$.

Lemma 5.2 *Let \mathcal{R} be \mathcal{T}_P or \mathcal{C}_P over Σ obtained from an instance P of PCP. Then \mathcal{R} is confluent (resp. terminating, right-most terminating, left-most terminating) if and only if w is confluent (resp. terminating, right-most terminating, left-most terminating) for every normal $w \in \Sigma^*$.*

Proof. Firstly we prove the termination part of the lemma. Since \Rightarrow -direction is trivial, consider \Leftarrow -direction.

Let $S_1 \xrightarrow{\mathcal{R}} S_2 \xrightarrow{\mathcal{R}} \dots$ be an infinite reduction sequence such that $\|S_1\|$ is minimal. We show a contradiction assuming $\|S_1\| > 0$. We have two cases that $S_1 = w\Xi_i S'$ and $S_1 = S'\Psi_i w$ for some normal w and some $S' \in \Sigma^*$.

- In the former case that $S_1 = w\Xi_i S'$, we can construct an infinite reduction sequence starting from at least one of w or $\Xi_i S'$ by applying Proposition 5.1(a) infinitely many times, which contradicts to the minimality of S_1 .
- In the latter case, we can show a contradiction in similar to the former case by using Proposition 5.1(c).

Secondly we prove the confluence part of the lemma. Since \Rightarrow -direction is trivial, consider \Leftarrow -direction. We show that every $S_1 \in \Sigma^*$ is confluent by induction on $\|S_1\|$. If $\|S_1\| = 0$, then S_1 is normal and it is confluent from the assumption. If $\|S_1\| > 0$, then we have two cases that $S_1 = w_1\Xi_i S'_1$ and $S_1 = S'_1\Psi_i w_1$ for some normal w_1 and some $S'_1 \in \Sigma^*$.

- In the former case, let $S_2 \xleftarrow{\mathcal{R}}^* w_1\Xi_i S'_1 \xrightarrow{\mathcal{R}}^* S_3$. By Proposition 5.1(b), we have $S_2 = w_2 S'_2$, $S_3 = w_3 S'_3$, $w_2 \xleftarrow{\mathcal{R}}^* w_1 \xrightarrow{\mathcal{R}}^* w_3$ and $S'_2 \xleftarrow{\mathcal{R}}^* \Xi_i S'_1 \xrightarrow{\mathcal{R}}^* S'_3$. Since w_1 is confluent from the assumption, we have $w_2 \downarrow_{\mathcal{R}} w_3$. Since $\Xi_i S'_1$ is confluent from the induction hypothesis, we have $S'_2 \downarrow_{\mathcal{R}} S'_3$. Therefore we have $S_2 = w_2 S'_2 \downarrow_{\mathcal{R}} w_3 S'_3 = S_3$.
- In the latter case, we can show it by using Proposition 5.1(d) in similar to the former case. \square

Note that this lemma is provable more elegantly by using a notion of persistency [17] in similar way to [4, 5]. However we gave the above proof to make the paper self-contained.

5.1 Termination analysis of \mathcal{T}_P

In the sequel, we analyze the termination property for \mathcal{T}_P .

We use a notation n^\diamond to represent either n or \tilde{n} for $n \in \{0, 1, 2\}$. We use a notation \vec{u} for $u_1 \dots u_k$.

Lemma 5.3 *Let P be an instance of PCP.*

(a) *If $u_1 \dots u_k \sim v_1 \dots v_k$ for some $(u_i, v_i) \in \overline{P}$, then $w \xrightarrow{\mathcal{T}_P}^+ w$ where*

$$w = \Xi_0(\tilde{0}0 \dots 0)_{\substack{u_1 \\ v_1}}^{\substack{u_1 \\ v_1}} \dots (\tilde{0}0 \dots 0)_{\substack{u_k \\ v_k}}^{\substack{u_k \\ v_k}} \Psi_0. \text{ Moreover, both of right-most reduction and left-most reduction are possible.}$$

(b) *If $\Xi_0 \chi \Psi_0 \xrightarrow{\mathcal{T}_P}^+ \Xi_0 \chi \Psi_0$ for some $\chi \in (\Sigma_c)^*$, then P has a solution.*

Proof. (a): We have a reduction sequence $\Xi_0(\tilde{0}0^\diamond \cdots 0^\diamond)_{\vec{u}, \vec{v}}^{\vec{u}} \Psi_0 \xrightarrow{\delta_2^*}$
 $\Xi_0(\tilde{0}0^\diamond \cdots 0^\diamond)_{\vec{u}, \vec{w}}^{\vec{u}} \Psi_0 \xrightarrow{\alpha_2 \cup \beta_2 \cup \gamma_2} \Xi_2(\tilde{2}2^\diamond \cdots 2^\diamond)_{\vec{u}, \vec{w}}^{\vec{u}} \Psi_2$. Here left-most reduction is possible. The right-most reduction exists by applying rules δ_2 as lazily as possible. Since $(u_i, v_i) \in \bar{P}$, we have a reduction sequence $\Xi_2(\tilde{2}2^\diamond \cdots 2^\diamond)_{\vec{u}, \vec{w}}^{\vec{u}} \Psi_2$
 $\xrightarrow{\gamma_1 \cup \beta_1 \cup \alpha_1} \Xi_0(\tilde{0}0^\diamond \cdots 0^\diamond)_{\vec{u}, \vec{w}}^{\vec{u}} \Psi_0$.
 (b): Let $\Xi_0 \chi \Psi_0 \xrightarrow{\mathcal{T}_P} \Xi_0 \chi \Psi_0$. From the construction of \mathcal{T}_P , a string $\Xi_2 \chi' \Psi_2$ must appear in this reduction sequence. From the reduction sequence $\Xi_0 \chi \Psi_0 \xrightarrow{\delta_2 \cup \alpha_2 \cup \beta_2 \cup \gamma_2} \Xi_2 \chi' \Psi_2$, the string χ must be in forms of $(\tilde{0}0 \cdots 0)_{\vec{u}'_1}^{u_1} \cdots (\tilde{0}0 \cdots 0)_{\vec{u}'_k}^{u_k}$ and χ' must be in forms of $(\tilde{2}2 \cdots 2)_{\vec{w}'_1}^{u_1} \cdots (\tilde{2}2 \cdots 2)_{\vec{w}'_k}^{u_k}$ where $\vec{u}' \sim \vec{v}'$. From the reduction sequence $\Xi_2 \chi' \Psi_2 = \Xi_2(\tilde{2}2^\diamond \cdots 2^\diamond)_{\vec{u}, \vec{w}}^{\vec{u}} \Psi_2$
 $\xrightarrow{\gamma_1 \cup \beta_1 \cup \alpha_1} \Xi_0(\tilde{0}0^\diamond \cdots 0^\diamond)_{\vec{u}', \vec{v}'}^{\vec{u}} \Psi_0 = \Xi_0 \chi \Psi_0$, we have $(u_i, v_i) \in \bar{P}$ for every i . Since \vec{u}' and \vec{v}' are copied from \vec{u} and \vec{v} respectively in the latter reduction sequence by β_1 -rules, we have $\vec{u}' = \vec{u}$ and $\vec{v}' = \vec{v}$. Thus we conclude $\vec{u} \sim \vec{v}$, which means that P has a solution. \square

Proof for Lemma 3.5

((i) \Rightarrow (ii) \wedge (iii)): By Lemma 5.3(a).

((ii) \vee (iii) \Rightarrow (iv)): Trivial.

((iv) \Rightarrow (i)): Let \mathcal{T}_P is not terminating. From Lemma 5.2, there is a non-terminating and normal string w . Infinite reduction sequences starting from w must contain a string starting with Ξ_0 and ending with Ψ_0 by the construction of \mathcal{T}_P . Thus the lemma follows from Lemma 5.3(b). \square

5.2 Confluence analysis of \mathcal{C}_P

In the sequel, we analyze the confluence property for \mathcal{C}_P .

The following propositions on the working area obtained from the construction of rules.

Proposition 5.4 *If $(\cdots)_{\vec{u}', \vec{v}'}^{\vec{u}} \xrightarrow{*} (\cdots)_{\vec{u}'', \vec{v}''}^{\vec{u}}$, then $u' \sim u''$ and $v' \sim v''$.*

Proposition 5.5 $\overset{*}{\leftarrow} \Theta = \overset{*}{\leftrightarrow} \Theta = \overset{*}{\rightarrow} \Theta$.

The following lemma shows that strings in a specific form are closed under reductions by Θ -rules. For example, $\{(222)_{\vec{u}_1}^{\vec{u}}, (022)_{\vec{u}_2}^{\vec{u}}, (002)_{\vec{u}_3}^{\vec{u}}\}$ is closed under the reductions.

Lemma 5.6 Let $\chi = \underbrace{(0^\diamond \cdots 0^\diamond)}_n \underbrace{p^\diamond p^\diamond \cdots p^\diamond}_{m'} \xrightarrow[\Theta]{*} \chi'$ where $m, n \geq 0$ and $p \in \{1, 2\}$. Then $\chi' = \underbrace{(0^\diamond \cdots 0^\diamond)}_{n'} \underbrace{p^\diamond p^\diamond \cdots p^\diamond}_{m'}$ for some $m', n' \geq 0$.

Proof. For any string in forms of χ for $p = 1$ (resp. $p = 2$), the only Θ_1 -rules (resp. Θ_2 -rules) are applicable, which produce a string in forms of χ' . \square

We state some properties on Θ_1 -rules.

Lemma 5.7 Consider the following strings for $i \leq j$:

$$\begin{aligned} \chi &= (\tilde{0}0 \cdots 0)_{u'_1}^{u_1} \cdots (\tilde{0}0 \cdots 0)_{u'_{i-1}}^{u_{i-1}} (\tilde{1}1 \cdots 1)_{u'_i}^{u_i} (\tilde{1}1 \cdots 1)_{u'_{i+1}}^{u_{i+1}} \cdots (\tilde{1}1 \cdots 1)_{u'_k}^{u_k}, \\ \chi' &= (\tilde{0}0 \cdots 0)_{u''_1}^{u_1} \cdots (\tilde{0}0 \cdots 0)_{u''_{j-1}}^{u_{j-1}} (\tilde{1}1 \cdots 1)_{u''_j}^{u_j} (\tilde{1}1 \cdots 1)_{u''_{j+1}}^{u_{j+1}} \cdots (\tilde{1}1 \cdots 1)_{u''_k}^{u_k}. \end{aligned}$$

If $\chi \xrightarrow[\Theta]{*} \chi'$ then $u_i = u'_i$, $v_i = v'_i$ and $(u_i, u'_i) \in \bar{P}$ for all $i \leq l < j$ and $u'_i = u''_i$ and $v'_i = v''_i$ for all $j \leq l$.

Proof. We have $\chi \xrightarrow[\Theta]{*} \chi'$ by Proposition 5.5. The lemma is proved by induction on the number of the rewrite steps. \square

Next we state some properties on Θ_2 -rules.

Lemma 5.8 Let $\chi = (\underline{2}^\diamond \underline{2}^\diamond \cdots \underline{2}^\diamond)_{u'_1}^u \xrightarrow[\Theta]{*} (0^\diamond \cdots 0^\diamond \underline{2}^\diamond)_{u''_1}^u$. Then $u'' \sim u' \sim v' \sim v''$.

Proof. We can prove, by induction on n , the claim that $\chi \xrightarrow[\Theta]{n} (0^\diamond \cdots 0^\diamond \underline{2}^\diamond)_{u'_1}^{u_1}$ implies $u'_1 \sim v'_1$. From this claim we have $u' \sim v'$. Hence the lemma follows from Proposition 5.4. \square

Lemma 5.9 If $w = \Xi_0(0^\diamond \cdots 0^\diamond)_{u'_1}^u \Psi_0 \xrightarrow[\mathcal{C}_P]{*} \Xi_0 \tilde{\underline{2}}_{x_3}^{x_1} \chi \Psi_2 = w'$ for some $\chi \in (\Sigma_c)^*$, then $u' \sim v'$.

Proof. Prove by induction on the number of rewrite steps in the reduction sequence. In the case that the first step is a reduction by α'_1 -rules, we have $w \xrightarrow[\alpha'_1]{*} \Xi_0 \chi' \Psi_1 \xrightarrow[\Theta_1]{*} \Xi_0 \chi'' \Psi_1 \xrightarrow[(\alpha'_1)^{-1}]{*} \Xi_0(0^\diamond \cdots 0^\diamond)_{u''_1}^u \Psi_0 \xrightarrow[\mathcal{C}_P]{*} w'$. The claim follows since $u' \sim u''$ and $v' \sim v''$ by Proposition 5.4 and $u'' \sim v''$ by the induction hypothesis.

Consider the case that the first step is a reduction by α_2 -rules. We have $w \xrightarrow[\alpha_2]{\Xi_0(0^\diamond \cdots 0^\diamond \underline{2})} w'$. If $(\alpha_2)^{-1}$ -rules are applied in the sequence, it is similar to the above case. Hence assume that $(\alpha_2)^{-1}$ -rules are not applied. Then, $w' = \Xi_0(\tilde{2}2^\diamond \cdots 2^\diamond) \Psi_2$ by Lemma 5.6. Thus $u' \sim v'$ follows from Proposition 5.5 and Lemma 5.8.

Consider the case that the first step is a reduction by δ_2 -rules. We have $w \xrightarrow[\delta_2]{\Xi_0(0^\diamond \cdots 0^\diamond)} w'$. The claim follows since $u' \sim u''$ and $v' \sim v''$ from Proposition 5.4 and $u'' \sim v''$ from the induction hypothesis. \square

Lemma 5.10 *If $w = \Xi_0(\tilde{0}0 \cdots 0) \Psi_0 \xrightarrow[\mathcal{C}_P]{*} \Xi_0(\tilde{1}x_1^{x_2}) \chi \Psi_1 = w'$ for some $\chi \in (\Sigma_c)^*$, then $u_1 \cdots u_k \sim u'_1 \cdots u'_k$, $v_1 \cdots v_k \sim v'_1 \cdots v'_k$ and $(u_i, v_i) \in \bar{P}$ for every i .*

Proof. Prove by induction on the number of rewrite steps in the reduction sequence. Consider the case that the first step is a reduction by α'_1 -rules and $(\alpha'_1)^{-1}$ -rules are not applied in the reduction. We have $w \xrightarrow[\alpha'_1]{*} w'' \xrightarrow[\mathcal{C}_P]{*} w'$, $u_k = u'_k$

and $v_k = v'_k$, where $w'' = \Xi_0(\tilde{0}0^\diamond \cdots 0^\diamond) \Psi_0 \xrightarrow[\mathcal{C}_P]{*} \Xi_0(\tilde{1}1 \cdots 1) \Psi_1$. Hence $w' = \Xi_0(\tilde{1}1 \cdots 1) \Psi_1$ by Lemma 5.6. By applying Lemma 5.7 with $i = 0$ and $j = k$ we obtain $u_l = u'_l$ and $v_l = v'_l$ for all $1 \leq l < k$ and $u'_k = u'_k$ and $v'_k = v'_k$. Hence we have $\vec{u} = \vec{u}'$ and $\vec{v} = \vec{v}'$. Since $\vec{u}' \sim \vec{u}''$ and $\vec{v}' \sim \vec{v}''$ by Proposition 5.4, $\vec{u} \sim \vec{u}'$ and $\vec{v} \sim \vec{v}'$ follow.

In the other cases, the proof is similar to that of Lemma 5.9. \square

Lemma 5.11 *Let P be an instance of PCP. If $w = \Xi_0(\tilde{1}x_1^{x_2}) \chi \Psi_1 \xrightarrow[\mathcal{C}_P]{*} \Xi_0(\tilde{2}x_2^{x_3}) \chi' \Psi_2 = w'$ for some $\chi, \chi' \in (\Sigma_c)^*$, then P has a solution.*

Proof. Let $w \xrightarrow[\mathcal{C}_P]{*} w'$. Then a string $\Xi_0 \chi'' \Psi_0$ must appear in this reduction and no underlined tag appears in χ'' from the construction of rules. Thus χ'' must be in forms of $\Xi_0(\tilde{0} \cdots 0) \Psi_0$; otherwise the underlined tag displayed in w do not move to next symbol of Ψ_i by Lemma 5.6 and the construction of rules. By Lemma 5.9 and Lemma 5.10, we have $\vec{u} \sim \vec{v}$ and $(u_i, v_i) \in \bar{P}$, which means P has a solution. \square

We need some more lemma in order to guarantee the confluence of \mathcal{C}_P when P has no solution.

Lemma 5.12 *Let w_1 and w_2 be normal strings over Σ^* . Then,*

- (a) $w_1 \xrightarrow[\mathcal{C}_P \setminus \gamma'_1]{*} w_2$ implies $w_1 \downarrow_{\mathcal{C}_P} w_2$, and
(b) $w_1 \xrightarrow[\mathcal{C}_P \setminus \gamma_2]{*} w_2$ implies $w_1 \downarrow_{\mathcal{C}_P} w_2$.

Proof. Before proving (a), we show the claim (*) that $w_1 \xleftarrow[\gamma_2]{*} w_2 \xrightarrow[\Theta]{*} w_3 \xrightarrow[\gamma_2]{*} w_4$ implies $w_1 \xrightarrow[\Theta]{*} w_4$ by induction on the number of rewrite steps. First of all w_2 must begin with $\Xi_0(\tilde{2})_{x_2}^{x_1}$ since it has a redex of γ_2 . Hence we can represent that $w_1 = \Xi_2(\tilde{2}X_1 \cdots X_n)_{v'}^u S'$, $w_2 = \Xi_0(\tilde{2}X_1 \cdots X_n)_{v'}^u S'$, $w_3 = \Xi_0(\tilde{2}X_1 \cdots X_n)_{v''}^u S''$ and $w_4 = \Xi_2(\tilde{2}X_1 \cdots X_n)_{v''}^u S''$ for $n \geq 0$, $X_i \in \{2, \tilde{2}\}$ and $S', S'' \in \Sigma^*$, where each tag of left-most symbol of S' and S'' is not 2 or $\tilde{2}$.

In the case that $S' = S'' = \Psi_2$, since $u' \sim u''$ and $v' \sim v''$ by Proposition 5.4, we have $w_1 \xrightarrow[\epsilon \cup \epsilon^{-1}]{*} w_4$. In the other cases, we can separate the reduction, from the construction of rules, into $S' \xrightarrow[\Theta]{*} S''$ and $w'_1 = \Xi_2(\tilde{2}X_1 \cdots X_n)_{v'}^u \xleftarrow[\gamma_2]{*} \Xi_0(\tilde{2}X_1 \cdots X_n)_{v'}^u \xrightarrow[\Theta]{*} \Xi_0(\tilde{2}X_1 \cdots X_n)_{v''}^u \xrightarrow[\gamma_2]{*} \Xi_2(\tilde{2}X_1 \cdots X_n)_{v''}^u = w'_4$. For the latter sequence, we have $w'_1 \xrightarrow[\epsilon \cup \epsilon^{-1}]{*} w'_4$ since $u' \sim u''$ and $v' \sim v''$ by Proposition 5.4. Therefore $w_1 \xrightarrow[\Theta]{*} w_4$.

Now we prove the lemma (a) by induction on the number k of reduction steps by γ_2 -rules in $w_1 \xrightarrow[\mathcal{C}_P \setminus \gamma'_1]{*} w_2$.

- ($k = 0$): It follows from Proposition 5.5.
- ($k = 1$): The reduction sequence can be represented as $w_1 \xrightarrow[\Theta]{*} w_3 \xleftrightarrow[\gamma_2]{*} w_4 \xrightarrow[\Theta]{*} w_2$. Then $w_1 \downarrow_{\mathcal{C}_P} w_2$ follows from Proposition 5.5.
- ($k > 1$): The reduction sequence can be represented as $w_1 \xrightarrow[\Theta]{*} w_3 \xleftrightarrow[\gamma_2]{*} w_4 \xrightarrow[\mathcal{C}_P \setminus \gamma'_1]{*} w_2$. If $w_3 \xrightarrow[\gamma_2]{*} w_4$ we have done by Proposition 5.5 and the induction hypothesis. Otherwise $w_1 \xrightarrow[\Theta]{*} w_3 \xleftarrow[\gamma_2]{*} w_4 \xrightarrow[\Theta]{*} w'_4 \xrightarrow[\gamma_2]{*} w'_2 \xrightarrow[\mathcal{C}_P \setminus \gamma'_1]{*} w_2$. Then $w_1 \downarrow_{\mathcal{C}_P} w_2$ by induction hypothesis since $w_1 \xrightarrow[\Theta]{*} w_3 \xrightarrow[\Theta]{*} w'_2 \xrightarrow[\mathcal{C}_P \setminus \gamma'_1]{*} w_2$ by the claim (*) above.

Before proving (b), we show the claim (**) that $w_1 \xleftarrow[\gamma'_1]{*} w_2 \xrightarrow[\Theta]{*} w_3 \xrightarrow[\gamma'_1]{*} w_4$ implies $w_1 \xrightarrow[\Theta]{*} w_4$ by induction on the number of rewrite steps. First of all w_2 must begin with $\Xi_0(\tilde{1})_{x_1}^{x_2}$ since it has a redex of γ'_1 . Hence we can represent that

$w_1 = \Xi_1(\tilde{1}X_1 \cdots X_n)_{v'}^u S'$, $w_2 = \Xi_0(\tilde{1}X_1 \cdots X_n)_{v'}^u S'$, $w_3 = \Xi_0(\tilde{1}X_1 \cdots X_n)_{v''}^u S''$ and $w_4 = \Xi_1(\tilde{1}X_1 \cdots X_n)_{v''}^u S''$ for $n \geq 0$, $X_i \in \{1, \tilde{1}\}$ and $S', S'' \in \Sigma^*$, where each tag of left-most symbol of S' and S'' is not 1 or $\tilde{1}$.

In the case that $S' = S'' = \Psi_1$, we have $u' = u''$ and $v' = v''$ by applying Lemma 5.7 with $i = j = 1$. Thus $w_1 = w_4$ follows. In the other cases, we can separate the reduction, from the construction of rules, into $S' \xrightarrow[\Theta]{*} S''$ and $w'_1 = \Xi_1(\tilde{1}X_1 \cdots X_n)_{v'}^u \xleftarrow[\gamma'_1]{*} \Xi_0(\tilde{1}X_1 \cdots X_n)_{v'}^u \xrightarrow[\Theta]{*} \Xi_0(\tilde{1}X_1 \cdots X_n)_{v''}^u \rightarrow \Xi_1(\tilde{1}X_1 \cdots X_n)_{v''}^u = w'_4$. For the latter sequence, we have $w'_1 = w'_4$ since $u' \sim u''$ and $v' \sim v''$ by Lemma 5.7. Therefore $w_1 \xrightarrow[\Theta]{*} w_4$.

By using the claim (**), the lemma (b) can be shown in similar to (a). \square

Proof for Lemma 4.3

Since \Rightarrow -direction is easy from the observation of Example 4.2, we show \Leftarrow -direction.

Assuming that P has no solution, let's show that \mathcal{C}_P is confluent. From Lemma 5.2, it is enough to consider $w_1 \xleftarrow[\mathcal{C}_P]{*} w_0 \xrightarrow[\mathcal{C}_P]{*} w_2$ for a normal string w_0 .

- Consider the case that w_0 starts with Ξ_0 and ends with Ψ_i for some $i \in \{0, 1, 2\}$. Assume that both of γ'_1 and γ_2 are applied in the reduction sequence. Then P must have a solution by Lemma 5.11, which is a contradiction. Hence at least one of γ'_1 or γ_2 rules cannot be applied in the reduction sequence.
- In either of following cases:
 - w_0 ends with Ψ_i for some $i \in \{0, 1, 2\}$ and all other symbols are of Σ_c ,
 - w_0 starts with Ξ_1 or Ξ_2 , and
 - w_0 starts with Ξ_0 and all other symbols are of Σ_c ,

It is easy to see that at least one of γ'_1 or γ_2 rules cannot be applied in the reduction sequence.

In any of the above cases, we have $w_1 \downarrow_{R_P} w_2$ by Lemma 5.12. \square

6 Length-two SRSs

Length-two SRSs are SRSs that consist of rules with length two, that is, $|l| = |r| = 2$ for every rule $l \rightarrow r$. In this section we give a transformation of a length preserving SRS over Σ_0 into a length-two SRS over Δ that preserves confluent property and termination property.

Let $\Sigma = \Sigma_0 \cup \{-\}$ and $m+1 (\geq 3)$ be the maximum length of rules in \mathcal{R} . Let $\Delta_0 = (\Sigma_0)^m$ and $\Delta = \Delta_0 \cup \{wv \mid w \in (\Sigma_0)^k, v \in \{-\}^{m-k}, 1 \leq k \leq m-1\}$.

The natural mapping $\phi : \Delta \rightarrow \Sigma^m$ is defined as $\phi(w) = w$. This mapping is naturally extended to $\phi : \Delta^* \rightarrow \Sigma^*$.

Example 6.1 Let $\Sigma_0 = \{a, b\}$ and $m = 2$. Then $\Delta_0 = \{aa, ab, ba, bb\}$, $\Delta = \Delta_0 \cup \{a-, b-\}$ and $\phi(ab \ bb \ a-) = abbba-$.

We give a transformation of a length preserving SRS \mathcal{R} into a length-two SRS $tw(\mathcal{R})$ over Δ .

$$tw(\mathcal{R}) = \{w_1w_2 \rightarrow w_3w_4 \mid w_i \in \Delta, \phi(w_1w_2) \xrightarrow{\mathcal{R}} \phi(w_3w_4)\}$$

Example 6.2 Let $\mathcal{R} = \{bbb \rightarrow aaa\}$ over $\Sigma_0 = \{a, b\}$. Then $tw(\mathcal{R})$ is the following length-two SRS over Δ , where Δ is displayed in Example 6.1.

$$tw(\mathcal{R}) = \begin{cases} bb \ b- \rightarrow aa \ a-, & bb \ ba \rightarrow aa \ aa, & bb \ bb \rightarrow aa \ ab, \\ ab \ bb \rightarrow aa \ aa, & bb \ bb \rightarrow ba \ aa \end{cases}$$

We say a string $w_1 \cdots w_n$ over Δ^* is *normal* if $w_1, \dots, w_{n-1} \in \Delta_0$. From the construction of $tw(\mathcal{R})$, all reachable strings from a normal string are also normal.

We define a mapping $\psi : \Delta^* \rightarrow (\Sigma_0)^*$ as $\psi(\alpha) = w$ where w is a string obtained from $\phi(\alpha)$ by removing all $-$'s. We define a mapping $\psi^{-1} : (\Sigma_0)^* \rightarrow \Delta^*$ as $\psi^{-1}(w) = \alpha$ where $\psi(\alpha) = w$ and α is normal. For example $\psi(ab \ bb \ a-) = abbba$ and $\psi^{-1}(abbba) = ab \ bb \ a-$. Trivially we have $\psi^{-1}(\psi(\alpha)) = \alpha$ for normal $\alpha \in \Delta^*$ and $\psi(\psi^{-1}(w)) = w$ for $w \in (\Sigma_0)^*$.

Proposition 6.3 (a) For a normal $\alpha_1 \in \Delta^*$, if $\alpha_1 \xrightarrow{tw(\mathcal{R})} \alpha_2$ then

$$\psi(\alpha_1) \xrightarrow{\mathcal{R}} \psi(\alpha_2)$$

(b) For $w_1 \in (\Sigma_0)^*$, if $w_1 \xrightarrow{\mathcal{R}} w_2$ then $\psi^{-1}(w_1) \xrightarrow{tw(\mathcal{R})} \psi^{-1}(w_2)$

Proof. From the construction of $tw(\mathcal{R})$. □

Lemma 6.4 Let \mathcal{R} an SRS. The SRS $tw(\mathcal{R})$ is confluent (resp. terminating, right-most terminating, left-most terminating) if and only if α is confluent (resp. terminating, right-most terminating, left-most terminating) for every normal $\alpha \in \Delta^*$.

Proof. We can prove it in similar to the proof of Lemma 5.2. Here $\Delta \setminus \Delta_0$ symbols play the same roles as Ψ_i symbols. □

Lemma 6.5 Let \mathcal{R} be an length preserving SRS. \mathcal{R} is terminating (resp. left-most terminating, right-most terminating) if and only if $tw(\mathcal{R})$ is terminating (resp. left-most terminating, right-most terminating).

Proof. (\Rightarrow): Let $tw(\mathcal{R})$ be non-terminating. By Lemma 6.4 we have an infinite reduction sequence for $tw(\mathcal{R})$ starting from a normal string. This direction follows from Proposition 6.3(a).

(\Leftarrow): Let \mathcal{R} be non-terminating. Then we have an infinite reduction sequence. By Proposition 6.3(b) we have an infinite reduction sequence for $tw(\mathcal{R})$.

This proof also works on either left-most case or right-most case. \square

Lemma 6.6 *Let \mathcal{R} be an length preserving SRS. \mathcal{R} is confluent if and only if $tw(\mathcal{R})$ is confluent.*

Proof. (\Rightarrow): Let $\beta_1 \xrightarrow{tw(\mathcal{R})}^* \alpha \xrightarrow{tw(\mathcal{R})}^* \beta_2$. We can assume that α is normal by Lemma 6.4. We have $\psi(\beta_1) \xrightarrow{\mathcal{R}}^* \psi(\alpha) \xrightarrow{\mathcal{R}}^* \psi(\beta_2)$ by Proposition 6.3(a). Since \mathcal{R} is confluent, there exists a string $w \in \Sigma_0^*$ such that $\psi(\beta_1) \xrightarrow{\mathcal{R}}^* w \xrightarrow{\mathcal{R}}^* \psi(\beta_2)$. Therefore we have $\beta_1 = \psi^{-1}(\psi(\beta_1)) \xrightarrow{tw(\mathcal{R})}^* \psi^{-1}(w) \xrightarrow{tw(\mathcal{R})}^* \psi^{-1}(\psi(\beta_2)) = \beta_2$ by Proposition 6.3(b).

(\Leftarrow): Let $u_1 \xrightarrow{\mathcal{R}}^* w \xrightarrow{\mathcal{R}}^* u_2$. We have $\psi^{-1}(u_1) \xrightarrow{tw(\mathcal{R})}^* \psi^{-1}(w) \xrightarrow{tw(\mathcal{R})}^* \psi^{-1}(u_2)$ by Proposition 6.3(b). Since \mathcal{R} is confluent, there exists a string $\alpha \in \Delta^*$ such that $\psi^{-1}(u_1) \xrightarrow{tw(\mathcal{R})}^* \alpha \xrightarrow{tw(\mathcal{R})}^* \psi^{-1}(u_2)$. Since α is normal, we have $u_1 = \psi(\psi^{-1}(u_1)) \xrightarrow{\mathcal{R}}^* \psi(\alpha) \xrightarrow{\mathcal{R}}^* \psi(\psi^{-1}(u_2)) = u_2$ by Proposition 6.3(a). \square

Theorem 6.7 *Confluence (termination, left-most termination, right-most termination) is an undecidable property for length-two SRSs.*

Proof. Directly obtained from Theorem 4.4 and Lemma 6.6 (Lemma 6.5). \square

Acknowledgment

We would like to thank the anonymous referees for their helpful comments and remarks. This work is partly supported by MEXT.KAKENHI #18500011 and #16300005.

References

- [1] F. Baader, T. Nipkow. *Term Rewriting and All That*, Cambridge University press, 1998.
- [2] A.-C. Caron. *Linear Bounded Automata and Rewrite Systems: Influence of Initial Configuration on Decision Properties*, Proc. of the Colloquium on Trees in Algebra and Programming(CAAP 91), LNCS, 493, pp.74–89, 1991.

- [3] N. Derchowitz. *Termination of Linear Rewriting Systems*, Proc. of the 8th International Colloquium on Automata, Languages and Programming (ICALP 18), LNCS, 115, pp.448–458, 1981.
- [4] A. Geser, A. Middeldorp, E. Ohlebusch, H. Zantema. *Relative Undecidability in Term Rewriting (Part 1: The Termination Hierarchy)*, Information and Computation, 178(1), pp.101–131, 2002.
- [5] A. Geser, A. Middeldorp, E. Ohlebusch, H. Zantema. *Relative Undecidability in Term Rewriting (Part 2: The confluence Hierarchy)*, Information and Computation, 178(1), pp.132–148, 2002.
- [6] G. Godoy, A. Tiwari, R. Verma. *On the Confluence of Linear Shallow Term Rewriting Systems*, Proc. of 20th Intl. Symposium on Theoretical Aspects of Computer Science (STACS 2003), LNCS, 2507, pp.85–96, 2003.
- [7] G. Godoy, A. Tiwari. *Confluence of Shallow Right-Linear Rewrite Systems*, Proc. of 14th Annual Conference on Computer Science Logic (CSL 2005), LNCS, 3634, pp.541–556, 2005.
- [8] G. Godoy, A. Tiwari. *Termination of Rewrite Systems with Shallow Right-Linear, Collapsing, and Right-Ground Rules*, Proc. of 20th International Conference on Automated Deduction (CADE 2005), LNCS, 3632, pp.164–176, 2005.
- [9] G. Huet, D. S. Lankford. *On the Uniform Halting Problem for Term Rewriting Systems*, Technical Report of INRIA, 283, 1978.
- [10] F. Jacquemard. *Reachability and Confluence are Undecidable for Flat Term Rewriting Systems*, Information Processing Letters 87(5), pp.265–270, 2003.
- [11] K. E. Knuth, P. B. Bendix. *Computational Problems in Abstract Algebra*, Pergamon Press, Oxford, pp.263–297, 1970.
- [12] I. Mitsuhashi, M. Oyamaguchi, Y. Ohta, and T. Yamada. *The Joinability and Related Decision Problems for Confluent Semi-Constructor TRSs*, Transactions of Information Processing Society of Japan, 47(5), pp.1502–1514, 2006.
- [13] I. Mitsuhashi, M. Oyamaguchi and F. Jacquemard. *The Confluence Problem for Flat TRSs*, Proc. of 8th Intl. Conf. on Artificial Intelligence and Symbolic Computation (AISC'06), LNAI 4120, pp.68–81, 2006.
- [14] M. Oyamaguchi, *The Church-Rosser Property for ground term rewriting systems is Decidable*, Theoretical Computer Science, 49, pp.43–79, 1987.
- [15] E. Post. *A Variant of a Recursively Unsolvable Problem*. Bulletin of the American Mathematical Society, 52, pp.264–268, 1946.

- [16] Y. Wang, M. Sakai, *Decidability of Termination for Semi-Constructor TRSs, Left-Linear Shallow TRSs and Related Systems*, Proc. of 17th International Conference on Term Rewriting and Applications (RTA 2006), LNCS, 4098, pp.343–356, 2006.
- [17] H. Zantema, *Termination of Term Rewriting: Interpretation and type elimination*, Journal of Symbolic Computation, 17, pp.23–50, 1994.