On Simulation-Completeness of Unraveling for Conditional Term Rewriting Systems*

Naoki Nishida[†], Masahiko Sakai and Toshiki Sakabe

Graduate School of Information Science, Nagoya University,
Furo-cho, Chikusa-ku, Nagoya 464–8603, Japan

December 27, 2005

Abstract

Transformations from conditional term rewriting systems (CTRSs) over (original) signatures into term rewriting systems (TRSs) over the extended signatures, are called unravelings. They are not always simulation-complete for any input CTRS. Here simulation-completeness means that every rewrite sequence of the unraveled CTRSs (the output TRSs), whose initial and final terms are over the original signatures, can be simulated by a rewrite sequence of the original CTRSs, and vice versa. We have proposed an unraveling which generates left-linear, right-linear and non-erasing TRSs from CTRSs satisfying some syntactic conditions, respectively. In this paper, we show two conditions that the unraveling is simulation-complete for CTRSs. One is that the unraveled CTRSs are right-linear and non-erasing, and the other is that they are left-linear. Under the latter condition, we assume that any redex introduced by extra variables is not reduced anywhere in the rewrite sequences of the unraveled CTRSs.

1 Introduction

Transformations from conditional term rewriting systems (CTRSs) into term rewriting systems (TRSs) are called *unravelings*. The unraveled CTRSs are approximations of the CTRSs, and they are used to analyze properties of the CTRSs, such as syntactic properties, termination and so on. The reason of such uses is that TRSs can be treated more easily than CTRSs. Marchiori has proposed some unravelings for several kinds of CTRSs to analyze *ultra-properties* and *modularity* of the CTRSs [4]. Ohlebusch has presented an unraveling for deterministic 3-CTRSs to analyze termination of logic programs [9]. To guarantee *effective termination* of CTRSs, termination of the unraveled CTRSs are required [4].

On the other hand, the unraveled CTRSs are not used to simulate rewrite sequences of the original CTRSs although reduction of TRSs are more simple

*This paper is a revised version of "IEICE Technical Report SS2004-18 (August 2004)".

[†]nishida@is.nagoya-u.ac.jp

than that of CTRSs. The reason is that unravelings are not always simulation-complete for any CTRS. Here simulation-completeness means that every rewrite sequence of the unraveled CTRSs (the output TRSs), whose initial and final terms are over the original signatures, can be simulated by a rewrite sequence of the original CTRSs, and vice versa.

An unraveling based on Ohlebusch's one has been proposed [6], and it has shown that the unraveling generates left-linear, right-linear and non-erasing TRSs from CTRSs satisfying some conditions, respectively.

In this paper, we show two conditions that the unraveling [6] is simulation-complete for CTRSs. One is that the unraveled CTRSs are right-linear and non-erasing, and the other is that they are left-linear. Under the latter condition, we assume that any redex introduced by extra variables is not reduced anywhere in the rewrite sequences of the unraveled CTRSs. We also show that under the first condition, neither right-linearity nor non-erasingness can be removed.

The unraveling which we treat in this paper, is used in an *inverse compiler* [6] that is a transformation from constructor TRSs into TRSs with extra variables (EV-TRSs) which compute inverses of functions in the input TRSs. In the first part of the compiler, constructor TRSs are transformed into deterministic 4-CTRSs, and then the CTRSs are transformed into EV-TRSs by using the unraveling. An example of the second transformation will be shown later (Example 2). It has been shown that all products of the inverse compiler are non-erasing, and those from left-linear (right-linear) TRSs are right-linear (left-linear, respectively). Hence, the results in this paper are useful for the inverse compiler. Since the first part of the inverse compiler generates 4-CTRSs and then the 4-CTRSs are unraveled to EV-TRSs, this paper treats 4-CTRSs and EV-TRSs.

2 Preliminaries

This paper follows general standard notations of term rewriting [1, 3, 8].

Let $\mathcal{V}$ be a countably infinite set of *variables*. The set of all *terms* over a *signature* $\mathcal{F}$ and $\mathcal{V}$ is denoted by $\mathcal{T}(\mathcal{F},\mathcal{V})$. Especially, we abbreviate the set $\mathcal{T}(\mathcal{F},\emptyset)$ of all *ground terms* to $\mathcal{T}(\mathcal{F})$. The set of variables occurring in either of terms $t_1,\ldots,t_n$ is represented by $\mathcal{V}ar(t_1,\ldots,t_n)$. *Identity* of terms is denoted by $\equiv$. The *root* symbol of a term t is represented by $\mathtt{root}(t)$. A term t is *linear* if no variable occurs twice in t.

The set of all *positions* of a term t is represented by $\mathcal{O}(t)$, and the sets of all *function symbol positions* and all *variable positions* of t are denoted by $\mathcal{O}_{\mathcal{F}}(t)$ and $\mathcal{O}_{\mathcal{V}}(t)$, respectively. The *root position* is represented by ε. For an n-hole *context* $C[,\ldots,]$ and terms $t_1,\ldots,t_n$, the notation $C[t_1,\ldots,t_n]_{p_1,\ldots,p_n}$ (or simply $C[t_1,\ldots,t_n]$) denotes the term which is obtained by replacing the occurrences of $\Box$ at positions $p_1,\ldots,p_n$ with $t_1,\ldots,t_n$ (from left to right). Given positions p and q of a term, we write $p \leq q$ if there exists p' such that $pp' = q$. Moreover, we write $p < q$ if $p' \neq \varepsilon$.

The *domain* and *range* of a *substitution* σ are denoted by $\mathcal{D}om(\sigma)$ and $\mathcal{R}an(\sigma)$, respectively. $\mathcal{S}ub(\mathcal{F},\mathcal{V})$ denotes the set of all substitutions whose range is over signature $\mathcal{F}$, that is, $\mathcal{S}ub(\mathcal{F},\mathcal{V}) = \{\sigma \mid \mathcal{R}an(\sigma) \subseteq \mathcal{T}(\mathcal{F},\mathcal{V})\}$. The *application* $\sigma(t)$ of σ to a term t may be abbreviated to $t\sigma$. The *composition* $\sigma\sigma'$ of substitutions σ and σ' is defined as $x\sigma\sigma' \equiv \sigma'(\sigma(x))$. When $\mathcal{D}om(\sigma) = \{x_1,\ldots,x_n\}$

and $\sigma(x_i) \equiv t_i$ for all $i \in \{1,\ldots,n\}$, we may write $\{ x_1 \mapsto t_i, \ldots, x_n \mapsto t_n \}$ for σ. For substitutions σ and σ', they are *equivalent*, written as $\sigma = \sigma'$, if $\mathcal{D}om(\sigma) = \mathcal{D}om(\sigma')$ and $\sigma(x) \equiv \sigma'(x)$ for all $x \in \mathcal{D}om(\sigma)$. The *restriction* $\sigma|_V$ of a substitution σ is defined as $\sigma|_V = \{ x \mapsto \sigma(x) \mid x \in \mathcal{D}om(\sigma) \cap V \}$. For substitutions σ and σ', we write $\sigma \leq \sigma'$ if there exists a substitution θ such that $\sigma\theta = \sigma'$.

An *oriented conditional rewrite rule* over a signature $\mathcal{F}$ is a triple (l, r, c), written as $l \to r \Leftarrow c$, where l is a non-variable term in $\mathcal{T}(\mathcal{F},\mathcal{V})$, r is a term in $\mathcal{T}(\mathcal{F},\mathcal{V})$, and c is a sequence $s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k$ of *oriented conditions* with $s_1, t_1, \ldots, s_k, t_k \in \mathcal{T}(\mathcal{F},\mathcal{V})$. The terms l and r are called the *left-hand side* (LHS) and the *right-hand side* (RHS) of the rule $l \to r \Leftarrow c$, respectively. The sequence c is called the *conditional part* of the rule. We say that the rule $l \to r \Leftarrow c$ is *unconditional* (or simply a *rewrite rule*), written as $l \to r$, if $k = 0$. The set of all variables in c is denoted by $\mathcal{V}ar(c)$. We may write the rule $l \to r \Leftarrow c$ as $\rho : l \to r \Leftarrow c$ with a unique label ρ. Variables which occur not in the LHS l but in either the RHS r or the conditional part c, are called *extra variables* of the rule ρ. The set of all extra variables of ρ is denoted by $\mathcal{EV}ar(\rho)$, that is, $\mathcal{EV}ar(\rho) = \mathcal{V}ar(r,c) \setminus \mathcal{V}ar(l)$ where $\mathcal{V}ar(r,c) = \mathcal{V}ar(r) \cup \mathcal{V}ar(c)$.

An *oriented conditional rewriting system* (oriented CTRS) over a signature $\mathcal{F}$ is a finite set of oriented conditional rewrite rules over $\mathcal{F}$. In this paper, we use the terminology "CTRS" as "oriented CTRS", and similarly for "conditional rewrite rule". Let R be a CTRS over signature $\mathcal{F}$. The *n-level rewrite relation* $\underset{n}{\to}_R$ of R is recursively defined as follows: $\underset{0}{\to}_R = \emptyset$, and $C[l\sigma]_p \underset{i+1}{\longrightarrow}_R C[r\sigma]_p$ if and only if $\rho : l \to r \Leftarrow s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k \in R$ and $s_j\sigma \overset{*}{\underset{i'}{\to}}_R t_j\sigma$ for every $j \in \{1,\ldots,k\}$ where $i' \leq i$ and $\overset{*}{\underset{i'}{\to}}_R$ is the reflexive and transitive closure of $\underset{i'}{\to}_R$. The *rewrite relation* $\to_R$ of the CTRS R is defined as $\to_R = \bigcup_{i \geq 0} \underset{i}{\to}_R$. To specify an applied rule ρ and a position p for $\to_R$, we write $\to_R^{[p,\rho]}$ or $\to_R^p$. We also write $\to_R^{q<}$ if $q < p$, and similarly for $q \leq$. A *rewrite sequence* of R is either a finite sequence $t_0 \to_R t_1 \to_R \cdots \to_R t_n$ or an infinite sequence $t_0 \to_R t_1 \to_R \cdots$, of terms in $\mathcal{T}(\mathcal{F},\mathcal{V})$. The term t_0 is called the *initial term* of the above rewrite sequence, and t_n is the *final term* when the sequence is finite. The *n-fold rewrite sequence* $t_0 \overset{n}{\to}_R t_n$ means that there exists a finite rewrite sequence $t_0 \to_R t_1 \to_R \cdots \to_R t_n$. The binary relation $\overset{*}{\to}_R$ is the reflexive and transitive closure of $\to_R$, $\overset{+}{\to}_R$ is the transitive closure of $\to_R$. Instances of all LHSs in R are called *redexes* (of R). A term t is called a *normal form* of R if there is no term u such that $t \to_R u$. The CTRS R is a *term rewriting system with extra variables* (EV-TRS) if every rule in R is unconditional, and R is a *term rewriting system* (TRS) if it is an EV-TRS and every rule $l \to r \in R$ satisfies $\mathcal{V}ar(l) \supseteq \mathcal{V}ar(r)$.

A conditional rewrite rule $\rho : l \to r \Leftarrow c$ will be classified according to the distribution of variables among l, r and c, as follows: ρ is in *type 1* if $\mathcal{V}ar(r,c) \subseteq \mathcal{V}ar(l)$, in *type 2* if $\mathcal{V}ar(r) \subseteq \mathcal{V}ar(l)$, in *type 3* if $\mathcal{V}ar(r) \subseteq \mathcal{V}ar(l,c)$, and in *type 4* if no restrictions. An *i-CTRS* contains only conditional rewrite rules in type i. A 1-CTRS R is called *normal* if for every rule $l \to r \Leftarrow s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k$ in R, $t_1, \ldots, t_k$ are normal forms of R.

Let R be a CTRS and $\rho : l \to r \Leftarrow c \in R$. The rule ρ is said to be *left-linear* (LL) if the LHS l of ρ is linear, it is *right-linear* (RL) if the RHS r of ρ is linear, and it is *non-erasing* (NE) if $\mathcal{V}ar(l) \subseteq \mathcal{V}ar(r)$. Let $c = s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k$. We say that the conditional rewrite rule ρ is *deterministic* if $\mathcal{V}ar(s_i) \subseteq$

$\mathcal{V}ar(l, t_1, \ldots, t_{i-1})$ for every $i \in \{1, \ldots, k\}$. The CTRS R satisfies a property $\mathcal{P}$ if every conditional rewrite rule in R satisfies $\mathcal{P}$.

Let R be a CTRS. A *parallel reduction* $\rightrightarrows_R$ of $\rightarrow_R$ is defined as, $C[s_1, \ldots, s_n]_{p_1,\ldots,p_n} \rightrightarrows_R C[t_1, \ldots, t_n]_{p_1,\ldots,p_n}$ if and only if $s_i \rightarrow_R^{\varepsilon} t_i$ for $1 \leq i \leq n$. We may write $\rightrightarrows_R^P$ for the above $\rightrightarrows_R$ where $P = \{p_1, \ldots, p_n\}$.

Basic reductions of EV-TRSs are defined similarly to that of TRSs [2, 5, 6]. Let R be an EV-TRS, and $\rho_i : l_i \rightarrow r_i \in R$ for every $i \in \{0, \ldots, n-1\}$. Let $t_1 \rightarrow_R^{[p_1,\rho_1]} t_2 \rightarrow_R^{[p_2,\rho_2]} \cdots \rightarrow_R^{[p_{n-1},\rho_{n-1}]} t_n$ be a rewrite sequence of R. We inductively define the sets of positions $B_1, \ldots, B_n$ from the sequence as follows:

- $B_1 \subseteq \mathcal{O}_{\mathcal{F}}(t_1)$, and
- $B_{i+1} = (B_i \setminus \{ q \in B_i \mid p_i \leq q \}) \cup \{ p_i q \mid q \in \mathcal{O}_{\mathcal{F}}(r_i) \}$ for $1 \leq i < n$

where B_1 is *closed under prefix* (that is, if $p < q$ and $q \in B_1$ then $p \in B_1$). For every $i \in \{1, \ldots, n\}$, positions in B_i are referred to as *basic positions* of t_i. We say that the above rewrite sequence is *based on* B_1 if $p_i \in B_i$ for $1 \leq i < n$. Especially, it is *basic* if $B_1 = \mathcal{O}_{\mathcal{F}}(t_1)$. Note that $B_2, \ldots, B_n$ in the above definition are clearly closed under prefix. To represent a rewrite sequence $s \xrightarrow{*}_R t$ based on a set B, we may write $B : s \xrightarrow[\mathrm{b}]{*}_R t$. Especially, we simply write $s \xrightarrow[\mathrm{b}]{*}_R t$ if $B = \mathcal{O}_{\mathcal{F}}(s)$.

3 Unraveling for Deterministic CTRSs

In this section, we introduce an unraveling for deterministic CTRSs, which generates left-linear, right-linear and non-erasing EV-TRSs from CTRSs satisfying some syntactic conditions, respectively [6].

For a set $A = \{a_1, \ldots, a_n\}$, $\vec{A}$ represents the unique list $a_1, \ldots, a_n$ of all elements a_is in A which is based on some ordering.

Definition 1 ([6]) *Let R be a deterministic CTRS over a signature $\mathcal{F}$. For every conditional rewrite rule $\rho : l \rightarrow r \Leftarrow s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k$ in R, we prepare k fresh function symbols $\mathsf{u}_1^\rho, \ldots, \mathsf{u}_k^\rho$ neither of which appears in $\mathcal{F}$. Then, the set $\mathbb{U}(\rho)$ of rewrite rules determined by ρ is defined as follows:*

$$\mathbb{U}(\rho) = \left\{ \begin{array}{rcl} l & \rightarrow & \mathsf{u}_1^\rho(s_1, \vec{X_1}), \\ \mathsf{u}_1^\rho(t_1, \vec{X_1}) & \rightarrow & \mathsf{u}_2^\rho(s_2, \vec{X_2}), \\ & \vdots & \\ \mathsf{u}_k^\rho(t_k, \vec{X_k}) & \rightarrow & r \end{array} \right.$$

where

$$X_i = \mathcal{V}ar(l, t_1, \ldots, t_{i-1}) \cap \mathcal{V}ar(r, t_i, s_{i+1}, t_{i+1}, \ldots, s_k, t_k).$$

Note that $\mathbb{U}(l' \rightarrow r') = \{ l' \rightarrow r' \}$. $\mathbb{U}$ is naturally extended over deterministic CTRSs: $\mathbb{U}(R) = \bigcup_{\rho \in R} \mathbb{U}(\rho)$.

Note that X_i is the set of variables that appear in either of $l, t_1, \ldots, t_{i-1}$ and also in either of $r, t_i, s_{i+1}, t_{i+1}, \ldots, s_k, t_k$. Relatively to the above definition of $\mathbb{U}$, we define the signature $\mathcal{F}_{\mathbb{U}}$ determined by $\mathcal{F}$ and R as $\mathcal{F} \cup \{ \mathsf{u}_i^\rho \mid \rho : l \rightarrow r \Leftarrow s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k \in R,\ 1 \leq i \leq k \}$. It is clear that the unraveled CTRS $\mathbb{U}(R)$ is over the signature $\mathcal{F}_{\mathbb{U}}$.

Example 2 Consider the following deterministic CTRS [6] where add^{-1} and mul^{-1} are function symbols that operates the inverses of addition and multiplication of natural numbers, and tp_2 is a constructor for tuples of two terms:

$$R_1 = \left\{ \begin{array}{l} add^{-1}(0) \to tp_2(0, y), \\ add^{-1}(s(z)) \to tp_2(s(x), y) \\ \qquad \Leftarrow add^{-1}(z) \twoheadrightarrow tp_2(x, y), \\ add^{-1}(add(x, y)) \to tp_2(x, y), \\ mul^{-1}(0) \to tp_2(x, 0), \quad mul^{-1}(0) \to tp_2(0, y), \\ mul^{-1}(s(z)) \to tp_2(s(x), s(y)) \\ \qquad \Leftarrow add^{-1}(z) \twoheadrightarrow tp_2(w, y), \\ \qquad\qquad mul^{-1}(w) \twoheadrightarrow tp_2(x, s(y)), \\ mul^{-1}(mul(x, y)) \to tp_2(x, y). \end{array} \right.$$

R_1 is unraveled by $\mathbb{U}$ as follows:

$$\mathbb{U}(R_1) = \left\{ \begin{array}{l} add^{-1}(0) \to tp_2(0, y), \\ add^{-1}(s(z)) \to \mathsf{u}_1(add^{-1}(z)), \\ \mathsf{u}_1(tp_2(x, y)) \to tp_2(s(x), y) \\ add^{-1}(add(x, y)) \to tp_2(x, y), \\ mul^{-1}(0) \to tp_2(x, 0), \\ mul^{-1}(0) \to tp_2(0, y), \\ mul^{-1}(s(z)) \to \mathsf{u}_2(add^{-1}(z)), \\ \mathsf{u}_2(tp_2(w, y)) \to \mathsf{u}_3(mul^{-1}(w), y), \\ \mathsf{u}_3(tp_2(x, s(y)), y) \to tp_2(s(x), s(y)), \\ mul^{-1}(mul(x, y)) \to tp_2(x, y). \end{array} \right.$$

The unraveling $\mathbb{U}$ we defined above is based on an unraveling for deterministic 3-CTRSs which has been proposed by Ohlebusch [9]. The difference between $\mathbb{U}$ and Ohlebusch's one is that $\vec{X_{i+1}}$ is used instead of $\vec{\mathcal{V}ar}(l), \vec{\mathcal{V}ar}(t_1), \ldots, \vec{\mathcal{V}ar}(t_i)$ in the definition of $\mathbb{U}(\rho)$. The variable list $\vec{\mathcal{V}ar}(l), \vec{\mathcal{V}ar}(t_1), \ldots, \vec{\mathcal{V}ar}(t_i)$ may contain variables not used in the later part $(s_{i+1}, t_{i+1}, \ldots, s_k, t_k, r)$, and such variables are redundant for simulations of rewrite sequences of CTRSs.

To analyze syntactic relations between a deterministic CTRS R and the EV-TRS $\mathbb{U}(R)$, we here define syntactic properties of deterministic CTRSs.

Definition 3 ([6]) *Let $\rho : l \to r \Leftarrow s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k$ be a conditional rewrite rule. The rule ρ is said to be*

- strictly non-erasing *if all variables in l appear in either of $r, s_1, t_1, \ldots, s_k, t_k$, and all variables in t_i appear in either of $r, t_i, s_{i+1}, t_{i+1}, \ldots, s_k, t_k$, that is,*
  - $\mathcal{V}ar(l) \subseteq \mathcal{V}ar(r, s_1, t_1, \ldots, s_k, t_k)$ *and*
  - $\mathcal{V}ar(t_i) \subseteq \mathcal{V}ar(s_{i+1}, t_{i+1}, \ldots, s_k, t_k, r)$ *for* $1 \leq i \leq k$,
- strictly left-linear *if $l, t_1, \ldots, t_k$ are linear and no variable in s_i appears in either of $r, t_i, s_{i+1}, t_{i+1}, \ldots, s_k, t_k$, that is, $\mathcal{V}ar(t_i) \cap \mathcal{V}ar(l, t_1, \ldots, t_{i-1}) = \emptyset$, and*
- strictly right-linear *if $r, s_1, \ldots, s_k$ are linear and no variable in s_i appears in either of $r, t_i, s_{i+1}, t_{i+1}, \ldots, s_k, t_k$, that is, $\mathcal{V}ar(s_i) \cap \mathcal{V}ar(r, t_i, s_{i+1}, t_{i+1}, \ldots, s_k, t_k) = \emptyset$.*

$\mathbb{U}$ enjoys the following nice properties associated with the above syntactic features.

Proposition 4 ([6]) *Let R be a deterministic CTRS.*

- *R is a 3-CTRS if and only if $\mathbb{U}(R)$ is a TRS.*
- *R is strictly NE if and only if $\mathbb{U}(R)$ is NE.*
- *R is strictly LL if and only if $\mathbb{U}(R)$ is LL.*
- *R is strictly RL if and only if $\mathbb{U}(R)$ is RL.*

Example 5 The CTRS R_1 in Example 2 is strictly RL and strictly NE, and the unraveled CTRS $\mathbb{U}(R_1)$ is RL and NE.

4 Completeness of Unraveling

Given an unraveling U, we say that U is *simulation-preserving* for a CTRS R over a signature $\mathcal{F}$ if for every terms s and t in $\mathcal{T}(\mathcal{F},\mathcal{V})$, $s \xrightarrow{*}_R t$ implies $s \xrightarrow{*}_{U(R)} t$. Conversely, U is *simulation-sound* for R over $\mathcal{F}$ if for every terms s and t in $\mathcal{T}(\mathcal{F},\mathcal{V})$, $s \xrightarrow{*}_{U(R)} t$ implies $s \xrightarrow{*}_R t$. Moreover, U is *simulation-complete* for R over $\mathcal{F}$ if for every terms s and t in $\mathcal{T}(\mathcal{F},\mathcal{V})$, $s \xrightarrow{*}_{U(R)} t$ coincides with $s \xrightarrow{*}_R t$.

Every unraveling is clearly simulation-preserving for all CTRSs in its target class (because simulation-preserving is a necessary condition for a transformation (from CTRSs to EV-TRSs) being an unraveling). However, unravelings are not always simulation-sound for any CTRS in their target classes, and hence simulation-incomplete. The following CTRS [8] is a counterexample for simulation-completeness of unravelings proposed by Marchiori [4] and Ohlebusch [9]. It also shows that our unraveling $\mathbb{U}$ is not simulation-complete in general.

Example 6 Consider the following deterministic CTRS:

$$R_2 = \begin{cases} f(x) \to x \Leftarrow x \twoheadrightarrow e, \\ h(x,x) \to g(x,x,f(k)), \quad g(d,x,x) \to A, \\ a \to c, \quad b \to c, \quad c \to e, \quad k \to l, \quad d \to m, \\ a \to d, \quad b \to d, \quad c \to l, \quad k \to m. \end{cases}$$

We obtain by $\mathbb{U}$ the following TRS:

$$\mathbb{U}(R_2) = \begin{cases} f(x) \to \mathsf{u}_4(x,x), \quad \mathsf{u}_4(e,x) \to x, \\ h(x,x) \to g(x,x,f(k)), \quad g(d,x,x) \to A, \\ a \to c, \quad b \to c, \quad c \to e, \quad k \to l, \quad d \to m, \\ a \to d, \quad b \to d, \quad c \to l, \quad k \to m. \end{cases}$$

There exists a rewrite sequence of $\mathbb{U}(R_2)$ from $h(f(a),f(b))$ to A, but there exists no rewrite sequence of R_2 from $h(f(a),f(b))$ to A, that is, $h(f(a),f(b)) \xrightarrow{*}_{\mathbb{U}(R_2)} A$ and $h(f(a),f(b)) \not\xrightarrow{*}_{R_2} A$. Therefore, $\mathbb{U}$ is not simulation-complete for R_2.

Now we discuss why $\mathbb{U}$ is not simulation-sound for R_2. We show a detail of $h(f(a), f(b)) \xrightarrow{*}_{\mathbb{U}(R_2)} A$ as follows:

$$\begin{array}{rl} h(f(a), f(b)) & \xrightarrow{*}_{\mathbb{U}(R_2)} h(\mathsf{u}_4(c,d), \mathsf{u}_4(c,d)) \\ & \to_{\mathbb{U}(R_2)} g(\mathsf{u}_4(c,d), \mathsf{u}_4(c,d), f(k)) \\ & \xrightarrow{*}_{\mathbb{U}(R_2)} g(d, \mathsf{u}_4(l,m), \mathsf{u}_4(l,m)) \to_{\mathbb{U}(R_2)} A. \end{array}$$

To construct the above reduction, the followings are required:

- to apply the rule $g(d, x, x) \to A$, the subterm $f(a)$ in the initial term is reduced to the term d,
- to apply the rule $h(x, x) \to g(x, x, f(k))$, the subterm $f(a)$ and $f(b)$ in the initial term are reduced to the same term, and
- to apply the rule $g(d, x, x) \to A$, the subterm $f(b)$ in the initial term and $f(k)$ are reduced to the same term.

The above requirements are summarized as, "$f(a)$, $f(b)$ and $f(k)$ are reduced to the same term d". In fact, this is however impossible on R_2. Nevertheless, $h(x, x) \to g(x, x, f(k))$ is used by reducing $f(a)$ and $f(b)$ to the same term $\mathsf{u}_4(c,d)$. After then, one of $\mathsf{u}_4(c,d)$ which comes from $f(a)$ is reduced to d, and the other $\mathsf{u}_4(c,d)$ from $f(b)$ is reduced to the term $\mathsf{u}_4(l,m)$ to be the same with $f(k)$. After then, $g(d, x, x) \to A$ is used.

These undesirable reductions may be caused by erasing rules and non-right-linear rules because $h(x, x) \to g(x, x, f(k))$ remains two deliver terms $\mathsf{u}_4(c,d)$ from $f(a)$ and $f(b)$ (each of which can performs a different role later), and $g(d, x, x) \to A$ erases two terms by considering them as the same (although it should not be done).

4.1 On Right-Linearity and Non-Erasingness

From the discussion before this subsection, $\mathbb{U}$ seems to be simulation-complete for a CTRS R if $\mathbb{U}(R)$ is RL and NE. In this subsection, we prove this claim.

The fact that every rewrite sequence of right-linear TRSs can be transformed to a basic rewrite sequence, have been proved by Middeldorp and Hamoen [5], and this result has been extended on EV-TRSs [6].

Proposition 7 ([6]) *Let R be a right-linear EV-TRS, s, t be terms, and θ be a substitution such that $\theta(x)$ is a normal form of R for all $x \in \mathcal{D}om(\theta)$. Then, $s\theta \xrightarrow{*}_R t$ implies $\mathcal{O}_{\mathcal{F}}(s) : s\theta \xrightarrow{*}_R t$.*

The above proposition says that every rewrite sequence of EV-TRSs can be transformed into a basic one. According to this fact, we can focus on basic rewrite sequences.

To treat reductions of strictly-LL CTRSs easily, we prepare a technical lemma. Let S, S' be sets, $\vec{A}$ be a list $a_1, \ldots, a_m$ on S, $\vec{B}$ be a list $b_1, \ldots, b_n$ on S', and $\to$ be a binary relation on $S \times S'$. Then, we write $\vec{A} \to \vec{B}$ if $m = n$ and $a_i \to b_i$ holds for all $i \in \{1, \ldots, n\}$. For a substitution σ, $\vec{A}\sigma$ denotes the unique list $a_1\sigma, \ldots, a_m\sigma$.

Lemma 8 *Let R be a CTRS, $l \to r \Leftarrow s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k \in R$ be strictly RL, and $\sigma_1, \ldots, \sigma_{k+1}$ be substitutions. Let $X_i = \mathcal{V}ar(l, t_1, \ldots, t_{i-1}) \cap \mathcal{V}ar(r, t_i, s_{i+1}, t_{i+1}, \ldots, s_k, t_k)$. If $s_i\sigma_i \xrightarrow{*}_R t_i\sigma_{i+1}$ and $\vec{X}_i\sigma_i \equiv \vec{X}_i\sigma_{i+1}$ for every $i \in \{1, \ldots, k\}$, then $l\sigma_1 \to_R r\sigma_{k+1}$.*

Proof. We only prove the case of $k = 1$. In this case, we have $X_1 = \mathcal{V}ar(l) \cap \mathcal{V}ar(r, t_1)$ ($\subseteq \mathcal{V}ar(l)$) from the assumption, and we also have $X_1 \cap \mathcal{V}ar(s_1) = \emptyset$ from the strictly-RL.

Let $V = \mathcal{EV}ar(\rho) \setminus \mathcal{V}ar(l, t_1)$ and $\sigma = \sigma_1|_{\mathcal{V}ar(l)} \cup \sigma_2|_{(\mathcal{V}ar(t_1) \setminus \mathcal{V}ar(l)) \cup V}$. It follows from the definition of extra variables that $V \cap \mathcal{V}ar(l, t_1) = \emptyset$. Hence σ is a substitution. We have $\vec{X}_1\sigma_1 \equiv \vec{X}_1\sigma_2$ from the assumptions. It follows from $\vec{X}_1\sigma_1 \equiv \vec{X}_1\sigma_2$, $X_1 = \mathcal{V}ar(l) \cap \mathcal{V}ar(r, t_1)$ and $X_1 \cap \mathcal{V}ar(s_1) = \emptyset$ that $\sigma|_{\mathcal{V}ar(t_1)} = (\sigma_1|_{\mathcal{V}ar(l)} \cup \sigma_2|_{\mathcal{V}ar(t_1) \setminus \mathcal{V}ar(l)}) = \sigma_2|_{\mathcal{V}ar(t_1)}$ and hence $t_1\sigma \equiv t_1\sigma_2$. Since $\mathcal{V}ar(s_1) \subseteq \mathcal{V}ar(l)$ from the definition of the determinism of the rule, we have $s_1\sigma \equiv s_1\sigma_1 \xrightarrow{*}_R t_1\sigma_2 \equiv t_1\sigma$, and hence $l\sigma_1 \equiv l\sigma \to_R r\sigma$.

Since $\mathcal{V}ar(r) \subseteq \mathcal{V}ar(t_1) \cup X_1 \cup V$ from the definitions of CTRSs and $\mathbb{U}$, and since $X_1 \subseteq \mathcal{V}ar(l)$ and $\vec{X}_1\sigma_1 \equiv \vec{X}_1\sigma_2$, we have $\sigma|_{X_1} = \sigma_1|_{X_1} = \sigma_2|_{X_1}$, $\sigma|_{\mathcal{V}ar(t_1) \setminus \mathcal{V}ar(l)} = \sigma_1|_{\mathcal{V}ar(t_1) \setminus \mathcal{V}ar(l)} = \sigma_2|_{\mathcal{V}ar(t_1) \setminus \mathcal{V}ar(l)}$ and $\sigma|_V = \sigma_2|_V$. Hence, $r\sigma \equiv r\sigma_2$. Therefore, we have $l\sigma_1 \to_R r\sigma_2$. □

To pick up parts of basic rewrite sequences, we show the following property related on rewrite sequences after reductions at the root position.

Proposition 9 *Let R be an EV-TRS, $l \to r \in R$, s, t be terms, $B \subseteq \mathcal{O}_{\mathcal{F}}(s)$, and θ be a substitution. Then, $B : s \xrightarrow[\mathrm{b}]{*}_R l\theta \xrightarrow[\mathrm{b}]{\varepsilon}_R r\theta \xrightarrow[\mathrm{b}]{*}_R t$ implies $B : s \xrightarrow[\mathrm{b}]{*}_R l\theta$, $\mathcal{O}_{\mathcal{F}}(l) : l\theta \xrightarrow[\mathrm{b}]{}_R r\theta$, and $\mathcal{O}_{\mathcal{F}}(r) : r\theta \xrightarrow[\mathrm{b}]{*}_R t$. For a context $C[\ ]_p$ and a set $B' \subseteq \mathcal{O}_{\mathcal{F}}(C[s]_p)$ such that $B = \{ q \mid pq \in B' \}$, $B' : C[s]_p \xrightarrow[\mathrm{b}]{*}{}^{p\leq}_R C[t]_p$ implies $B : s \xrightarrow[\mathrm{b}]{*}_R t$.*

From the definition of $\xrightarrow[\mathrm{b}]{}_R$, this proposition holds clearly.

The following lemma is a generalized one of the claim that $\mathbb{U}$ is simulation-sound (with respect to basic sequences) for every strictly-RL and strictly-NE CTRS.

Lemma 10 *Let R be a strictly-RL and strictly-NE CTRS over a signature $\mathcal{F}$. Let $s, t \in \mathcal{T}(\mathcal{F}, \mathcal{V})$, $B \subseteq \mathcal{O}_{\mathcal{F}}(s)$, and $\theta \in \mathcal{S}ub(\mathcal{F}, \mathcal{V})$. Then, $B : s \xrightarrow[\mathrm{b}]{*}_{\mathbb{U}(R)} t$ implies $s \xrightarrow{*}_R t$.*

Proof. For a sequence $B : s \xrightarrow[\mathrm{b}]{n}_{\mathbb{U}(R)} t$, we prove this lemma by induction on the lexicographic combination of n and the structure of the term s.

The case that s is a variable, holds obviously since $B : s \xrightarrow[\mathrm{b}]{n}_{\mathbb{U}(R)} t$ implies $s \equiv t$.

The remaining case is that s is not a variable. We first consider the subcase that the sequence $s \xrightarrow{n}_{\mathbb{U}(R)} t$ does not contain any reduction at the root position, that is, $s \xrightarrow{n}{}^{\varepsilon<}_{\mathbb{U}(R)} t$. This subcase is easily proved by the induction hypothesis because the sequence can be decomposed to rewrite sequences of which numbers of steps are less equal than n, and which start from the proper subterms of s.

We consider the remaining subcase. From the construction of $\mathbb{U}(R)$, $\mathbb{U}(R)$ has the unique rule for every $\mathsf{u}_i^\rho \notin \mathcal{F}$. Now the root symbols of s and t are in $\mathcal{F}$ ($\mathtt{root}(s), \mathtt{root}(t) \neq \mathsf{u}_j^\rho$). According to these facts, we can assume that

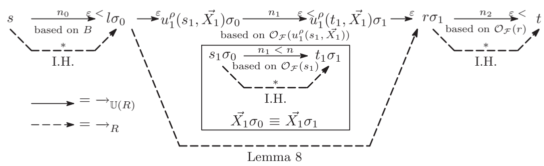

Figure 1: Proof sketch of Lemma 10.

- $\rho : l \to r \Leftarrow s_1 \twoheadrightarrow t_1 \in R$,
- $l \to \mathsf{u}_1^{\rho}(s_1, \vec{X_1})$, $\mathsf{u}_1^{\rho}(t_1, \vec{X_1}) \to r \in \mathbb{U}(R)$ where $X_1 = \mathcal{V}ar(l) \cap \mathcal{V}ar(r, t_1)$, and
- $B : s \xrightarrow[\mathrm{b}]{n_0}{}^{\varepsilon<}_{\mathbb{U}(R)} l\sigma_0 \xrightarrow[\mathrm{b}]{}{}^{\varepsilon}_{\mathbb{U}(R)} \mathsf{u}_1^{\rho}(s_1, \vec{X_1})\sigma_0 \xrightarrow[\mathrm{b}]{n_1}{}^{\varepsilon<}_{\mathbb{U}(R)} \mathsf{u}_1^{\rho}(t_1, \vec{X_1})\sigma_1 \xrightarrow[\mathrm{b}]{}{}^{\varepsilon}_{\mathbb{U}(R)} r\sigma_1 \xrightarrow[\mathrm{b}]{n_2}{}_{\mathbb{U}(R)} t$ where $n_0 + n_1 + n_2 = n - 2$ (see Fig. 1).

From Proposition 9, we have $\mathcal{O}_{\mathcal{F}}(r) : r\sigma_1 \xrightarrow[\mathrm{b}]{n_2}{}_{\mathbb{U}(R)} t$. Since $t \in \mathcal{T}(\mathcal{F}, \mathcal{V})$ and $\mathbb{U}(R)$ is NE by Proposition 4, any term introduced by $\sigma_1|_{\mathcal{V}ar(r)}$ does not contain any symbol in $\mathcal{F}_{\mathbb{U}} \setminus \mathcal{F}$, that is, $\mathcal{R}an(\sigma_1|_{\mathcal{V}ar(r)}) \subseteq \mathcal{T}(\mathcal{F}, \mathcal{V})$. We also obtain $\mathcal{R}an(\sigma_0) \subseteq \mathcal{T}(\mathcal{F}, \mathcal{V})$ from the same reason above for the part $\mathsf{u}_1^{\rho}(s_1, \vec{X_1})\sigma_0 \xrightarrow[\mathrm{b}]{n_1}{}^{\varepsilon<}_{\mathbb{U}(R)} \mathsf{u}_1^{\rho}(t_1, \vec{X_1})\sigma_1$. Hence, we have $l\sigma_0, s_1\sigma_0, t_1\sigma_1, r\sigma_1 \in \mathcal{T}(\mathcal{F}, \mathcal{V})$. We have $\mathcal{O}_{\mathcal{F}}(s_1) : s_1\sigma_0 \xrightarrow[\mathrm{b}]{n_1}{}_{\mathbb{U}(R)} t_1\sigma_1$ from Proposition 9, and we also have $\vec{X_1}\sigma_0 \equiv \vec{X_1}\sigma_1$ from the basicness of that sequence. By the induction hypothesis, we have $s \xrightarrow{*}_R l\sigma_0$, $s_1\sigma_0 \xrightarrow{*}_R t_1\sigma_1$ and $r\sigma_1 \xrightarrow{*}_R t$. From Lemma 8, we have $l\sigma_0 \to_R r\sigma_1$. Therefore, we obtain $s \xrightarrow{*}_R l\sigma_0 \to_R r\sigma_1 \xrightarrow{*}_R t$ (see Fig. 1). □

Finally, we obtain the following result.

Theorem 11 $\mathbb{U}$ *is simulation-complete for every strictly-RL and strictly-NE CTRS.*

Proof. Let $s, t \in \mathcal{T}(\mathcal{F}, \mathcal{V})$, and suppose that $s \xrightarrow{*}_{\mathbb{U}(R)} t$. Since $\mathbb{U}(R)$ is RL by Proposition 4, we have a basic rewrite sequence $s \xrightarrow[\mathrm{b}]{*}{}_{\mathbb{U}(R)} t$ from Proposition 7. Then, from Lemma 10, we obtain $s \xrightarrow{*}_R t$. □

The conjecture in [6] that $\mathbb{U}$ is simulation-complete for every strictly-NE CTRS R, does not hold. Consider the CTRS R_3 obtained from R_2 by replacing A with x. Although we have $h(f(a), f(b)) \xrightarrow{*}_{\mathbb{U}(R_3)} l$, there is no rewrite sequence of R_3 from $h(f(a), f(b))$ to l, that is, $h(f(a), f(b)) \not\xrightarrow{*}_{R_3} l$. Therefore, there exists a strictly-NE CTRS for which $\mathbb{U}$ is not simulation-complete.

Similarly, the other condition "strictly-NE" cannot be also removed. Consider the strictly-RL and non-NE CTRS $R_4 = \{ f(x) \to e \Leftarrow d \twoheadrightarrow l, h(x, x) \to A \}$. This CTRS is unraveled to $\mathbb{U}(R_4) = \{ f(x) \to \mathsf{u}_5(d), \mathsf{u}_5(l) \to e, h(x, x) \to A \}$. Though we have $h(f(a), f(b)) \xrightarrow{*}_{\mathbb{U}(R_4)} A$, there is no rewrite sequence of R_4 from $h(f(a), f(b))$ to A, that is, $h(f(a), f(b)) \not\xrightarrow{*}_{R_4} A$. Therefore, there exists a strictly-RL CTRS for which $\mathbb{U}$ is not simulation-complete.

4.2 On Left-Linearity

Marchiori has shown that his unraveling [4] for normal CTRSs are simulation-complete for CTRSs if the unraveled CTRSs are left-linear. In addition, the left-linearity seems to be another solution of the problems as we have shown in Example 2. Hence, our unraveling $\mathbb{U}$ seems to be simulation-complete for a strictly-LL CTRS R ($\mathbb{U}(R)$ is left-linear from Proposition 4). This claim holds for 3-CTRSs but not for 4-CTRSs. The reason is that extra variables appearing only in RHSs make mischief by introducing garbage terms. We can solve this problem if we restrict rewrite sequences to *EV-safe* ones in which any redex introduced by extra variables is not reduced anywhere. In this subsection, we prove the above claim for such sequences.

We first define *EV-safe* rewrite sequences of EV-TRSs [6, 7]. In these sequences, any redex introduced by extra variables are not reduced anywhere. Let R be an EV-TRS. EV-safe rewrite sequences are defined similarly to basic rewrite sequences, by adding $\{\, p_i q' r \mid p_i q \in B_i, q \in \mathcal{O}_{\mathcal{V}}(l_i), l_i|_q \equiv r_i|_{q'} \,\}$ to B_{i+1} in the definition of $\underset{\text{b}}{\rightarrow}_R$. We say that the above rewrite sequence is *EV-safe with respect to* B_1 if $p_i \in B_i$ for $1 \leq i < n$. Especially, it is *EV-safe* if $B_1 = \mathcal{O}_{\mathcal{F}}(t_1)$. To represent the above rewrite sequence, we may write $B_1 : t_1 \overset{*}{\underset{\text{ev}}{\rightarrow}}_R t$. Especially, we simply write $t_1 \overset{*}{\underset{\text{ev}}{\rightarrow}}_R t_n$ if $B_1 = \mathcal{O}_{\mathcal{F}}(t_1)$. The definition of $\underset{\text{ev}}{\rightarrow}_R$ is different from that in [6, 7], however they are equivalent essentially. A typical instance of EV-safe rewrite sequences is a rewrite sequence obtained by substituting a normal form for each extra variable.

We again prepare a technical lemma to treat reductions of strictly-LL CTRSs easily.

Lemma 12 *Let R be a CTRS, $l \to r \Leftarrow s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k \in R$ be strictly LL, and $\sigma_1, \ldots, \sigma_{k+1}$, be substitutions. Let $X_i = \mathcal{V}ar(l, t_1, \ldots, t_{i-1}) \cap \mathcal{V}ar(r, t_i, s_{i+1}, t_{i+1}, \ldots, s_k, t_k)$. If $s_i\sigma_i \overset{*}{\rightarrow}_R t_i\sigma_{i+1}$ and $\vec{X}_i\sigma_i \overset{*}{\rightarrow}_R \vec{X}_i\sigma_{i+1}$ for every $i \in \{1, \ldots, k\}$, then $l\sigma_1 \overset{+}{\rightarrow}_R r\sigma_{k+1}$.*

Proof. We only prove the case of $k = 1$. In this case, we have $X_1 = \mathcal{V}ar(l) \cap \mathcal{V}ar(r, t_1)$ ($\subseteq \mathcal{V}ar(l)$) from the assumption, and we also have $X_1 \cap \mathcal{V}ar(t_1) = \emptyset$ from the strictly-LL.

Let $V = \mathcal{EV}ar(\rho) \setminus \mathcal{V}ar(l, t_1)$ and $\sigma = \sigma_1|_{\mathcal{V}ar(l)} \cup \sigma_2|_{\mathcal{V}ar(t_1) \cup V}$. It follows from the definition of extra variables that $V \cap \mathcal{V}ar(l, t_1) = \emptyset$. Since $X_1 \cap \mathcal{V}ar(t_1) = \emptyset$, we have $\mathcal{V}ar(l) \cap \mathcal{V}ar(t_1) = \emptyset$. Then, σ is a substitution. Since $\mathcal{V}ar(s_1) \subseteq \mathcal{V}ar(l)$ from the determinism of the rule, we have $s_1\sigma \equiv s_1\sigma_1 \overset{*}{\rightarrow}_R t_1\sigma_2 \equiv t_1\sigma$, and hence $l\sigma_1 \equiv l\sigma \rightarrow_R r\sigma$.

From the definitions of CTRSs and $\mathbb{U}$, $\mathcal{V}ar(r) \subseteq \mathcal{V}ar(t_1) \cup X_1 \cup V$. From the assumption, $X_1 \subseteq \mathcal{V}ar(l)$. Hence we have $r\sigma \equiv r(\sigma|_{\mathcal{V}ar(l)} \cup \sigma|_{\mathcal{V}ar(t_1)} \cup \sigma|_V)$ $\equiv r(\sigma_1|_{\mathcal{V}ar(l)} \cup \sigma_2|_{\mathcal{V}ar(t_1)} \cup \sigma_2|_V)$. Since $\vec{X}_1\sigma_1 \overset{*}{\rightarrow}_R \vec{X}_1\sigma_2$ and $X_1 \subseteq \mathcal{V}ar(l)$, we have $r\sigma \overset{*}{\rightarrow}_R r\sigma_2$. Therefore, $l\sigma_1 \overset{+}{\rightarrow}_R r\sigma_2$. □

Let R be a CTRS over a signature $\mathcal{F}$, $\rho_i : l_i \to r_i \in \mathbb{U}(R)$ for $i \geq 1$, and $T \subseteq \mathcal{T}(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$. Suppose that a rewrite sequence $t_0 \underset{\text{ev}}{\rightarrow}^{[p_1,\rho_1]}_{\mathbb{U}(R)} t_1 \underset{\text{ev}}{\rightarrow}^{[p_2,\rho_2]}_{\mathbb{U}(R)} \cdots$ satisfies $t_0 \equiv C'_0[l_1\sigma_1]_{p_1}$ and $t_i \equiv C_i[r_i\sigma_i]_{p_i} \equiv C'_i[l_{i+1}\sigma_{i+1}]_{p_{i+1}}$ for $i \geq 1$. This rewrite sequence is said to be *EV-arranged on* T if $\mathcal{R}an(\sigma_i|_{\mathcal{EV}ar(\rho_i)}) \subseteq T$ for every $i \geq 1$. EV-safe rewrite sequences of the unraveled CTRSs has the following property associated with the EV-arranged property.

Proposition 13 *Let R be a strictly LL CTRS over a signature $\mathcal{F}$, and $s, t \in \mathcal{T}(\mathcal{F}, \mathcal{V})$. If $s \xrightarrow{*}_{\mathbb{U}(R)} t$ is EV-safe, then it is EV-arranged on $\mathcal{T}(\mathcal{F}, \mathcal{V})$.*

Proof. *(Sketch)* We can prove by induction on steps n that for $s \in \mathcal{T}(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$, $t \in \mathcal{T}(\mathcal{F}, \mathcal{V})$ and $\theta \in \mathcal{S}ub(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$, if $\mathcal{O}_{\mathcal{F}}(s) : s\theta \xrightarrow[\text{ev}]{n}_{\mathbb{U}(R)} t$, then there exists a substitution $\theta' \in \mathcal{S}ub(\mathcal{F}, \mathcal{V})$ such that $s\theta' \xrightarrow{n}_{\mathbb{U}(R)} t$ and $\theta' \leq \theta$. From this lemma, we can shown the claim of this proposition.

As another proof, we can prove this proposition by using the completeness of narrowing of EV-TRSs. As the completeness of narrowing, it holds that for a ground term $s' \in \mathcal{T}(F)$ and a term $t' \in \mathcal{T}(\mathcal{F}, \mathcal{V})$, if $s' \xrightarrow[\text{ev}]{n}_{\mathbb{U}(R)} t'$, then there exist a term u and substitutions θ, σ such that $s' \overset{n}{\leadsto}{}^{\sigma}_{\mathbb{U}(R)} u$ and $u\theta \equiv t'$ [6, 7]. Here we can assume that $\mathcal{D}om(\sigma)$ only contains extra variables which are of rules applied at each step. Since $t' \in \mathcal{T}(\mathcal{F}, \mathcal{V})$, we have $u \in \mathcal{T}(\mathcal{F}, \mathcal{V})$ and we can suppose that $\mathcal{R}an(\theta) \subseteq \mathcal{T}(\mathcal{F}, \mathcal{V})$. From the construction of $\mathbb{U}$, every rule $l \to r$ in $\mathbb{U}(R)$ satisfies that all proper subterms of l are in $\mathcal{T}(\mathcal{F}, \mathcal{V})$. Since σ is determined by most general unifiers of terms and LHSs at each step and l is linear from the LL of $\mathbb{U}(R)$, we have $\mathcal{R}an(\sigma) \subseteq \mathcal{T}(\mathcal{F}, \mathcal{V})$. Hence, we have $\mathcal{R}an(\sigma\theta) \subseteq \mathcal{T}(\mathcal{F}, \mathcal{V})$. This means that terms in $\mathcal{T}(\mathcal{F}, \mathcal{V})$ are substituted for each extra variables in every step of $s' \xrightarrow[\text{ev}]{n}_{\mathbb{U}(R)} t'$. Then, $s' \xrightarrow[\text{ev}]{n}_{\mathbb{U}(R)} t'$ is EV-arranged on $\mathcal{T}(\mathcal{F}, \mathcal{V})$. In every rewrite sequence, variables in redexes can be considered as constants. Therefore, the above claim also holds for $s \xrightarrow{*}_{\mathbb{U}(R)} t$. □

The following lemma is a generalized one of the main claim in this subsection.

Lemma 14 *Let R be a strictly LL CTRS, $s \in \mathcal{T}(\mathcal{F}, \mathcal{V})$, t be a linear term in $\mathcal{T}(\mathcal{F}, \mathcal{V})$, and $\theta \in \mathcal{S}ub(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$. If $s \overset{n}{\rightrightarrows}_{\mathbb{U}(R)} t\theta$ (for some $n \geq 0$) is EV-arranged on $\mathcal{T}(\mathcal{F}, \mathcal{V})$, then there exists a substitution $\theta' \in \mathcal{S}ub(\mathcal{F}, \mathcal{V})$ such that $s \xrightarrow{*}_R t\theta'$ and $t\theta' \overset{n'}{\rightrightarrows}_{\mathbb{U}(R)} t\theta$ for some $n' \leq n$.*

Proof. We prove this lemma by induction on the lexicographic combination of n and the structure of the term s.

Since the case that $n = 0$ or $s \in \mathcal{V}$ holds clearly, we concentrate the remaining case, that is, $n > 0$ and $s \notin \mathcal{V}$.

Consider the subcase that $s \overset{n}{\rightrightarrows}_R t\theta$ does not contain any reduction at the root position, that is, $s \overset{n}{\rightrightarrows}{}^{\varepsilon<}_{\mathbb{U}(R)} t\theta$ where $\rightrightarrows^{\varepsilon<}_R = \bigcup_{P \cap \{\varepsilon\} = \emptyset} \rightrightarrows^P_{\mathbb{U}(R)}$. This subcase is easily proved by the induction hypothesis because the sequence can be decomposed to rewrite sequences of which numbers of steps are less equal than n, and which start from the proper subterms of s.

Consider the remaining subcase. From the construction of $\mathbb{U}(R)$, $\mathbb{U}(R)$ has the unique rule for every $\mathsf{u}^\rho_i \notin \mathcal{F}$. Now the root symbols of s and t are in $\mathcal{F}$ ($\mathtt{root}(s), \mathtt{root}(t) \neq \mathsf{u}^\rho_j$). According to these facts, we can assume that

- $\rho : l \to r \Leftarrow s_1 \twoheadrightarrow t_1 \in R$,
- $l \to \mathsf{u}^\rho_1(s_1, \vec{X_1})$, $\mathsf{u}^\rho_1(t_1, \vec{X_1}) \to r \in \mathbb{U}(R)$ where $X_1 = \mathcal{V}ar(l) \cap \mathcal{V}ar(r, t_1)$, and
- $s \overset{n_0}{\rightrightarrows}{}^{\varepsilon<}_{\mathbb{U}(R)} l\sigma_0 \rightrightarrows^{\{\varepsilon\}}_{\mathbb{U}(R)} \mathsf{u}^\rho_1(s_1, \vec{X_1})\sigma_0 \overset{n_1}{\rightrightarrows}{}^{\varepsilon<}_{\mathbb{U}(R)} \mathsf{u}^\rho_1(t_1, \vec{X_1})\sigma_1 \rightrightarrows^{\{\varepsilon\}}_{\mathbb{U}(R)} r\sigma_1 \overset{n_2}{\rightrightarrows}_{\mathbb{U}(R)} t\theta$ where $n_0 + n_1 + n_2 = n - 2$ (see Fig. 2).

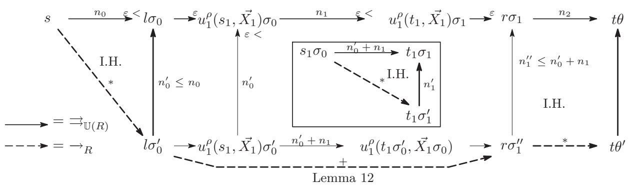

Figure 2: Proof sketch of Lemma 14.

It is clear that $l \in \mathcal{T}(\mathcal{F},\mathcal{V})$ and $n_0 < n$. Then, by the induction hypothesis, there exists a substitution $\sigma'_0 \in \mathcal{S}ub(\mathcal{F},\mathcal{V})$ such that $s \xrightarrow{*}_R l\sigma'_0$ and $l\sigma'_0 \stackrel{n'_0}{\rightrightarrows}_{\mathbb{U}(R)} l\sigma$ for $n'_0 \leq n_0$. Then, it follows from $\mathcal{V}ar(s_1) \cup X_1 \subseteq \mathcal{V}ar(l)$ that we have $s_1\sigma'_0 \stackrel{n'_0+n_1}{\rightrightarrows}_{\mathbb{U}(R)} t_1\sigma_1$ and $\vec{X_1}\sigma'_0 \stackrel{n'_0+n_1}{\rightrightarrows}_{\mathbb{U}(R)} \vec{X_1}\sigma_1$. It is clear that $s_1\sigma'_0, t_1 \in \mathcal{T}(\mathcal{F},\mathcal{V})$ and $n'_0+n_1 < n$. Then, by the induction hypothesis, there exists a substitution $\sigma'_1 \in \mathcal{S}ub(\mathcal{F},\mathcal{V})$ such that $s_1\sigma'_0 \xrightarrow{*}_R t_1\sigma'_1$ and $t_1\sigma'_1 \stackrel{n'_1}{\rightrightarrows}_{\mathbb{U}(R)} t_1\sigma_1$ for $n'_1 \leq n_1$.

Let $V = \mathcal{EV}ar(\rho) \setminus \mathcal{V}ar(l,t_1)$ and $\sigma''_1 = \sigma'_1|_{\mathcal{V}ar(t_1)} \cup \{\, x \mapsto \sigma_0(x) \mid x \in X_1 \,\} \cup \sigma_1|_V$. Since R is strictly LL, we have $\mathcal{V}ar(t_1) \cap X_1 = \emptyset$. We also have $V \cap (\mathcal{V}ar(t_1) \cup X_1) = \emptyset$. from the definition of extra variables. Then, σ''_1 is clearly a substitution. Since the sequence is EV-arranged on $\mathcal{T}(\mathcal{F},\mathcal{V})$, we can assume that $\mathcal{R}an(\sigma_1|_V) \subseteq \mathcal{T}(\mathcal{F},\mathcal{V})$. Hence $\sigma''_1 \in \mathcal{S}ub(\mathcal{F},\mathcal{V})$. Now we have $t_1\sigma''_1 \equiv t_1\sigma'_1$, $t_1\sigma'_1 \stackrel{n'_1}{\rightrightarrows}_{\mathbb{U}(R)} t_1\sigma_1$ and $\vec{X_1}\sigma''_1 \equiv \vec{X_1}\sigma'_0 \stackrel{n'_0+n_1}{\rightrightarrows}_{\mathbb{U}(R)} \vec{X_1}\sigma_1$. It follows from $\mathcal{V}ar(r) \subseteq \mathcal{V}ar(t_1) \cup X_1 \cup V$ that $r\sigma''_1 \stackrel{n''_1}{\rightrightarrows}_{\mathbb{U}(R)} r\sigma_1$ for $n''_1 \leq n'_0+n_1$. Then, we have $r\sigma''_1 \stackrel{n''_1}{\rightrightarrows}_{\mathbb{U}(R)} r\sigma_1 \stackrel{n_2}{\rightrightarrows}_{\mathbb{U}(R)} t\theta$. It is clear that $r\sigma''_1 \in \mathcal{T}(\mathcal{F},\mathcal{V})$ and $n''_1 + n_2 < n$. By the induction hypothesis, there exists a substitution $\theta' \in \mathcal{S}ub(\mathcal{F},\mathcal{V})$ such that $r\sigma''_1 \xrightarrow{*}_R t\theta'$ and $t\theta' \stackrel{n'}{\rightrightarrows}_{\mathbb{U}(R)} t\theta$ for $n' \leq n$.

Now we have $s_1\sigma'_0 \xrightarrow{*}_R t_1\sigma'_1 \equiv t_1\sigma''_1$ and $\vec{X_1}\sigma'_0 \equiv \vec{X_1}\sigma''_1$. It follows from Lemma 12 that $l\sigma_0 \xrightarrow{+}_R r\sigma''_1$. Therefore, we obtain $s \xrightarrow{*}_R l\sigma'_0 \xrightarrow{+}_R r\sigma''_1 \xrightarrow{*}_R t\theta'$ (see Fig. 2). □

Theorem 15 *For every strictly-LL CTRS R, $\mathbb{U}$ is simulation-complete with respect to* $\underset{\text{ev}}{\rightarrow}_{\mathbb{U}(R)}$.

Proof. Let $s,t \in \mathcal{T}(\mathcal{F},\mathcal{V})$, and suppose that $s \xrightarrow{*}_{\mathbb{U}(R)} t$ is EV-safe. Since $s \xrightarrow{*}_{\mathbb{U}(R)} t$ is EV-safe, we can assume that $s \xrightarrow{*}_{\mathbb{U}(R)} t$ is EV-arranged on $\mathcal{T}(\mathcal{F},\mathcal{V})$ from Proposition 13. Since it is known that $\rightrightarrows_{R'} \subseteq \xrightarrow{*}_{R'} \subseteq \stackrel{*}{\rightrightarrows}_{R'}$ for every (EV-)TRS R', $s \xrightarrow{*}_{\mathbb{U}(R)} t$ implies $s \stackrel{*}{\rightrightarrows}_{\mathbb{U}(R)} t$, and then $s \stackrel{*}{\rightrightarrows}_{\mathbb{U}(R)} t$ implies $s \xrightarrow{*}_R t$ from Lemma 14. Therefore, this theorem holds. □

It is clear that for every TRS R, all rewrite sequences of R are EV-safe. Hence we obtain the following corollary from Theorem 15.

Corollary 16 $\mathbb{U}$ *is simulation-complete for every strictly-LL 3-CTRS.*

Here we show a counterexample that the above theorem does not hold for non-EV-safe sequences. Consider the CTRS $R_5 = \{\, e \to f(x) \Leftarrow l \twoheadrightarrow d,\; A \to h(x,x) \,\}$ and the EV-TRS $\mathbb{U}(R_5) = \{\, e \to \mathsf{u}_6(l),\; \mathsf{u}_6(d) \to f(x),\; A \to h(x,x) \,\}$. We have $A \xrightarrow{*}_{\mathbb{U}(R_5)} h(f(a), f(b))$ but $A \overset{*}{\not\to}_{R_5} h(f(a), f(b))$.

Remark that the results in this paper (Theorem 11 and 15) do not hold for Ohlebusch's unraveling (on which $\mathbb{U}$ is based) because Proposition 4 does not hold for his unraveling.

5 Conclusion

We have shown that the unraveling $\mathbb{U}$ is simulation-complete for every deterministic CTRS which is either strictly-LL or strictly-RL and strictly-NE, where rewrite sequences of the unraveled CTRSs are EV-safe under the former condition. From the discussion why simulation-completeness does not hold, $\mathbb{U}$ seems to be simulation-complete with respect to EV-safe rewrite sequences for a CTRS R if every rule in $\mathbb{U}(R)$ is either LL or RL and NE, in other words, if every rule in R is either strictly LL or strictly RL and strictly NE. As a future work, we will prove this conjecture (or show a counterexample). We are also interested in a relationship between LL, RL, NE and the existence of extra variables.

References

[1] F. Baader and T. Nipkow: "Term Rewriting and All That", Cambridge University Press (1998).

[2] J.-M. Hullot: "Canonical forms and unification", Proceedings of the 5th International Conference on Automated Deduction, Vol. 87 of Lecture Notes in Computer Science, Springer-Verlag, pp. 318–334 (1980).

[3] J. W. Klop: "Term rewriting systems", Handbook of Logic in Computer Science (Eds. by S. Abramsky, D. M. Gabbay and T. S. E. Maibaum), Vol. 2, Oxford University Press, New York, pp. 2–116 (1992).

[4] M. Marchiori: "Unravelings and ultra-properties", Proceedings of the 5th International Conference on Algebraic and Logic Programming, Vol. 1139 of Lecture Notes in Computer Science, Springer-Verlag, pp. 107–121 (1996).

[5] A. Middeldorp and E. Hamoen: "Completeness results for basic narrowing", Applicable Algebra in Engineering, Communication and Computing, **5**, pp. 213–253 (1994).

[6] N. Nishida: "Transformational Approach to Inverse Computation in Term Rewriting", PhD thesis, Nagoya University (2004).

[7] N. Nishida, M. Sakai and T. Sakabe: "Narrowing-based simulation of term rewriting systems with extra variables and its termination proof", Functional and Constraint Logic Programming (Ed. by G. Vidal), Vol. 86 of Electronic Notes in Theoretical Computer Science, Issue 3, Elsevier, (18 pages) (2003).

[8] E. Ohlebusch: "Advanced Topics in Term Rewriting", Springer-Verlag (2002).

[9] E. Ohlebusch: "Termination of logic programs: Transformational methods revisited", Applicable Algebra in Engineering, Communication and Computing, **12**, 1-2, pp. 73–116 (2001).

A Lemma for the proof of Proposition 13

We here give a lemma to prove Proposition 13 along the first part of the proofsketch.

We inductively define a new operation $\sqcap$ on a term and a linear term which are unifiable, as follows:

- $s \sqcap x = (x, \{x \mapsto s\})$,
- $x \sqcap t = (t, \emptyset)$ where $t \notin \mathcal{V}$, and
- $f(s_1, \ldots, s_n) \sqcap f(t_1, \ldots, t_n) = (f(u_1, \ldots, u_n), \sigma_1 \cup \cdots \cup \sigma_n)$ where $s_i \sqcap t_i = (u_i, \sigma)$ for every $i \in \{1, \ldots, n\}$.

Since the second argument of $\sqcap$ is linear, it is clear that $\sigma_1 \cup \cdots \cup \sigma_n$ is a substitution. It is also clear that $s \sqcap t = (u, \sigma)$ implies $s \equiv u\sigma$, $\mathcal{D}om(\sigma) \subseteq \mathcal{V}ar(t)$ and $\mathcal{VR}an(\sigma) \subseteq \mathcal{V}ar(s)$. The result (u, σ) of $s \sqcap t$ means that s is divided to u and σ (that is, $s \equiv u\sigma$), and that u is a maximal term of the common part of s and t.

Lemma 17 *Let R be a strictly LL CTRS over a signature $\mathcal{F}$, $s \in \mathcal{T}(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$, $t \in \mathcal{T}(\mathcal{F}, \mathcal{V})$ and let $\theta \in \mathcal{S}ub(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$. If $\mathcal{O}_{\mathcal{F}}(s) : s\theta \xrightarrow[\mathrm{ev}]{n}_{\mathbb{U}(R)} t$, then there exists a substitution $\theta' \in \mathcal{S}ub(\mathcal{F}, \mathcal{V})$ such that $\theta' \leq \theta$ and $s\theta'\theta'' \xrightarrow{n}_{\mathbb{U}(R)} t$ for any substitution $\theta'' \in \mathcal{S}ub(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$.*

Proof. We can prove this lemma by induction on steps n.

Since the case of $n = 0$ is obvious, we only consider the remaining case, that is, $n > 0$. From the assumption of the lemma, we suppose that

- $\mathcal{D}om(\theta) \subseteq \mathcal{V}ar(s)$,
- $s \equiv C[v]$,
- $v \notin \mathcal{V}$,
- $l \to r \in R$,
- $\mathcal{V}ar(s) \cap \mathcal{V}ar(l, r) = \emptyset$,
- $v\theta \equiv l\sigma$,
- $\mathcal{D}om(\sigma) \subseteq \mathcal{V}ar(l, r)$, and
- $\mathcal{O}_{\mathcal{F}}(s) : s\theta \equiv C\theta[l\sigma] \xrightarrow[\mathrm{ev}]{}_{\mathbb{U}(R)} C\theta[r\sigma] \xrightarrow[\mathrm{ev}]{n-1}_{\mathbb{U}(R)} t$.

Since v and l are unifiable and l is linear, there exist a term $v_0 \in \mathcal{T}(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$ and a substitution $\theta_0 \in \mathcal{S}ub(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$ such that $v \sqcap l = (v_0, \theta_0)$, and hence $v_0\theta_0 \equiv v$. Let $\sigma' = \sigma|_{\mathcal{D}om(\sigma) \setminus \mathcal{D}om(\theta_0)}$. Since $\mathcal{V}ar(s) \cap \mathcal{V}ar(l, r) = \emptyset$, $\mathcal{D}om(\sigma) \subseteq \mathcal{V}ar(l, r)$, $\mathcal{VR}an(\theta_0) \subseteq \mathcal{V}ar(v)$ and $v\theta \equiv l\sigma$, we have $\sigma = (\theta_0\theta) \cup \sigma|_{\mathcal{D}om(\sigma) \setminus \mathcal{D}om(\theta_0)}$, $\theta_0\sigma'$

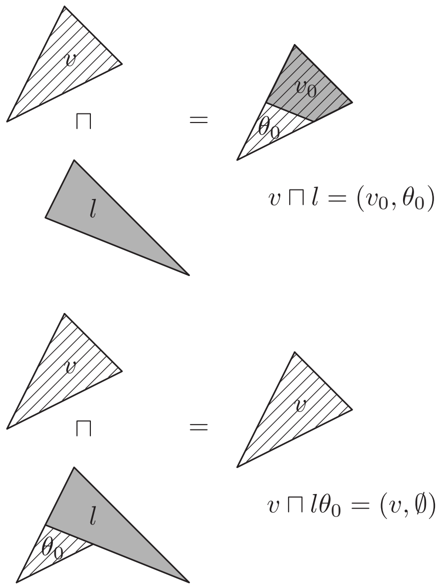

Figure 3: The relationship between terms and substitution in $v \sqcap l = (v_0, \theta_0)$.

$=_{\sigma'} C\theta[r\sigma] \equiv (C[r\theta_0])(\theta \cup \sigma')$. It follows from the definition of EV-safety that $\mathcal{O}_{\mathcal{F}}(C[r\theta_0]) : (C[r\theta_0])(\theta \cup \sigma') \xrightarrow[\mathrm{ev}]{n-1}_{\mathbb{U}(R)} t$. By the induction hypothesis, there exists a substitution $\eta \in \mathcal{S}ub(\mathcal{F}, \mathcal{V})$ such that $\eta \leq \theta \cup \sigma'$ and $(C[r\theta_0])\eta\theta' \xrightarrow{n-1}_{\mathbb{U}(R)} t$ for any substitution $\theta' \in \mathcal{S}ub(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$.

Since $v(\theta \cup \sigma') \equiv v\theta \equiv l\sigma \equiv l(\theta_0\theta \cup \sigma') \equiv l(\theta_0\theta \cup \theta_0\sigma') \equiv l\theta_0(\theta \cup \sigma')$, there exists a unifier η' of v and $l\theta_0$ such that $\eta' \leq \theta \cup \sigma'$. Then, it follows from $v \sqcap l = (v_0, \theta_0)$ that $l\theta_0 \sqcap l = (l, \theta_0)$ (see Fig. 3). From the construction of $\mathbb{U}(R)$, every proper subterm of l is in $\mathcal{T}(\mathcal{F}, \mathcal{V})$. Hence we can assume that $\eta' \in \mathcal{S}ub(\mathcal{F}, \mathcal{V})$. Now we suppose that $\eta \leq \eta' \leq \theta \cup \sigma'$. Let a substitution η'' such that $\eta' = \eta\eta''$. Then, $s\eta' \equiv C\eta'[v\eta'] \equiv C\eta'[l\theta_0\eta'] \to_{\mathbb{U}(R)} C\eta'[r\theta_0\eta'] \equiv (C[r\theta_0])\eta' \equiv (C[r\theta_0])\eta\eta''$. Since $(C[r\theta_0])\eta\theta' \xrightarrow{n-1}_{\mathbb{U}(R)} t$ for any substitution $\theta' \in \mathcal{S}ub(\mathcal{F}_{\mathbb{U}}, \mathcal{V})$, we have $(C[r\theta_0])\eta\eta''\theta' \xrightarrow{n-1}_{\mathbb{U}(R)} t$ for any substitution θ'. Hence, we obtain $s\eta'\theta' \to_{\mathbb{U}(R)} (C[r\theta_0])\eta'\theta' \xrightarrow{n-1}_{\mathbb{U}(R)} t$ for any substitution θ'. It is clear that $\eta'|_{\mathcal{V}ar(s)} \leq \theta$. Therefore, we have $\eta'|_{\mathcal{V}ar(s)} \leq \theta$ and $s\eta'|_{\mathcal{V}ar(s)}\theta' \equiv s\eta'\theta' \xrightarrow{n}_{\mathbb{U}(R)} t$ for any substitution θ'. $\square$