On Transforming Narrowing Trees into Regular Tree Grammars Generating Ranges of Substitutions*

Naoki Nishida

Graduate School of Informatics Nagoya University Nagoya, Japan nishida@i.nagoya-u.ac.jp Yuya Maeda

Graduate School of Informatics Nagoya University Nagoya, Japan yuya@trs.css.i.nagoya-u.ac.jp

The grammar representation of a narrowing tree for a syntactically deterministic conditional term rewriting system and a pair of terms is a regular tree grammar that generates expressions for substitutions obtained by all possible innermost-narrowing derivations that start with the pair and end with non-narrowable terms. In this paper, under a certain syntactic condition, we show a transformation of the grammar representation of a narrowing tree into another regular tree grammar that generates the ranges of substitutions generated by the grammar representation. In our previous work, such a transformation is restricted to the ranges w.r.t. a given single variable, and thus, the usefulness is limited. We extend the previous transformation by representing the range of a substitution as a tuple of terms, which is obtained by the coding for finite trees.

1 Introduction

Conditional term rewriting [22, Chapter 7] is known to be more complicated than unconditional term rewriting in the sense of analyzing properties, e.g., *operational termination* [15], *confluence* [25], and *reachability* [5]. A popular approach to the analysis of conditional term rewriting systems (CTRSs, for short) is to transform a CTRS into an unconditional term rewriting system (a TRS, for short) that is in general an overapproximation of the CTRS in terms of reduction. This approach enables us to use existing techniques for the analysis of TRSs. For example, a CTRS is operationally terminating if the *unraveled* TRS [16, 22] is terminating [4]. To prove termination of the unraveled TRS, we can use many techniques for proving termination of TRSs (cf. [22]). On the other hand, it is not so easy to analyze *reachability* that is relevant to, e.g., *(in)feasibility* of conditions.

Let us consider to prove confluence of the following *syntactically deterministic* 3-CTRS [22, Example 7.1.5] defining the *gcd* operator over the natural numbers represented by 0 and s:

$$\mathcal{R}_1 = \left\{ \begin{array}{ll} x < 0 \rightarrow \mathsf{false}, & 0 - \mathsf{s}(y) \rightarrow 0, \\ 0 < \mathsf{s}(y) \rightarrow \mathsf{true}, & x - 0 \rightarrow x, \\ \mathsf{s}(x) < \mathsf{s}(y) \rightarrow x < y, & \mathsf{s}(x) - \mathsf{s}(y) \rightarrow x - y, \\ \mathsf{gcd}(x, x) \rightarrow x, & \\ \mathsf{gcd}(\mathsf{s}(x), 0) \rightarrow \mathsf{s}(x), & \mathsf{gcd}(\mathsf{s}(x), \mathsf{s}(y)) \rightarrow \mathsf{gcd}(x - y, \mathsf{s}(y)) \Leftarrow y < x \twoheadrightarrow \mathsf{true}, \\ \mathsf{gcd}(0, \mathsf{s}(y)) \rightarrow \mathsf{s}(y), & \mathsf{gcd}(\mathsf{s}(x), \mathsf{s}(y)) \rightarrow \mathsf{gcd}(\mathsf{s}(x), y - x) \Leftarrow x < y \twoheadrightarrow \mathsf{true} \end{array} \right\}$$

A transformational approach in [9, 8] does not succeed in proving confluence of \mathcal{R}_1 . On the other hand, a direct approach to reachability analysis to prove *infeasibility* of the *conditional critical pairs* (i.e., non-existence of substitutions satisfying conditions), which is implemented in some confluence provers, does

^{*}This work was partially supported by JSPS KAKENHI Grant Number JP17H01722.

not prove confluence of \mathcal{R}_1 well, either. For example, \mathcal{R}_1 has the following six critical pairs:

$$\begin{array}{l} \langle & \mathsf{s}(x), \; \mathsf{gcd}(x-x,\mathsf{s}(x)) \; \rangle \Leftarrow x < x \twoheadrightarrow \mathsf{true}, \\ \langle \; \mathsf{gcd}(x-x,\mathsf{s}(x)), \; \mathsf{s}(x) & \rangle \Leftarrow x < x \twoheadrightarrow \mathsf{true}, \\ \langle \; \; \mathsf{s}(x), \; \mathsf{gcd}(\mathsf{s}(x),x-x) \; \rangle \Leftarrow x < x \twoheadrightarrow \mathsf{true}, \\ \langle \; \mathsf{gcd}(\mathsf{s}(x),x-x), \; \mathsf{s}(x) & \rangle \Leftarrow x < x \twoheadrightarrow \mathsf{true}, \\ \langle \; \mathsf{gcd}(x-y,\mathsf{s}(y)), \; \mathsf{gcd}(\mathsf{s}(x),y-x) \; \rangle \Leftarrow x < y \twoheadrightarrow \mathsf{true}, \; y < x \twoheadrightarrow \mathsf{true}, \\ \langle \; \mathsf{gcd}(\mathsf{s}(x),y-x), \; \mathsf{gcd}(x-y,\mathsf{s}(y)) \; \rangle \Leftarrow x < y \twoheadrightarrow \mathsf{true}, \; y < x \twoheadrightarrow \mathsf{true}, \end{array}$$

An operationally terminating CTRS is confluent if all critical pairs of the CTRS are infeasible (cf. [1, 3]). Operational termination of \mathcal{R}_1 can be proved by, e.g., AProVE [6]. To prove infeasibility of the critical pairs above, it suffices to show both (a) non-existence of terms t such that $t < t \rightarrow_{\mathcal{R}_1}^*$ true, and (b) non-existence of terms t_1, t_2 such that $t_1 < t_2 \rightarrow_{\mathcal{R}_1}^*$ true and $t_2 < t_1 \rightarrow_{\mathcal{R}_1}^*$ true. Thanks to the meaning of <, it would be easy for a human to notice that such terms t, t_1, t_2 do not exist. However, it is not so easy to mechanize a way to show non-existence of t, t_1, t_2 . In fact, confluence provers for CTRSs, ConCon [24], CO3 [18], and CoScart [7], based on e.g., transformations of CTRSs into TRSs or reachability analysis for infeasibility of conditional critical pairs failed to prove confluence of \mathcal{R}_1 (see Confluence Competition 2016¹ and 2017,² 327.trs). In addition, a *semantic approach* in [14, 13] cannot prove confluence of \mathcal{R}_1 using AGES [10], a tool for generating logical models of order-sorted first-order theories—non-existence of t_1, t_2 above cannot be proved via its web interface with default parameters.

The non-existence of a term t with $t < t \rightarrow_{\mathcal{R}_1}^*$ true can be reduced to the non-existence of substitutions θ such that $x < x \rightsquigarrow_{\theta,\mathcal{R}_1}^*$ true, where \rightsquigarrow denotes the *narrowing* step [12]. In addition, the non-existence of such substitutions can be reduced to the emptiness of the set of such substitutions, i.e., the emptiness of $\{\theta \mid x < x \rightsquigarrow_{\theta,\mathcal{R}_1}^*$ true}. From this viewpoint, for pairs of terms, the enumeration of substitutions obtained by narrowing would be useful in analyzing rewriting that starts with instances of the pairs. To analyze sets of substitutions derived by *innermost* narrowing, *narrowing trees* [20] are useful. For example, infeasibility of conditional critical pairs of a normal 1-CTRS can be proved by using the *grammar representation* of a narrowing tree [19].

In this paper, under a certain syntactic condition, we show a transformation of the grammar representation of a narrowing tree into a *regular tree grammar* [2] (an RTG, for short) that generates the ranges of substitutions generated by the grammar representation.

Let \mathcal{R} be a syntactically deterministic 3-CTRS (a 3-SDCTRS, for short) that is a constructor system, s a basic term, and t a constructor term, where basic terms are of the form $f(u_1, \ldots, u_n)$ with a defined symbol f and constructor terms u_1, \ldots, u_n . A narrowing tree [20, 19] of \mathcal{R} with the root pair $s \twoheadrightarrow t$ is a finite representation that defines the set of substitutions θ such that the pair $s \twoheadrightarrow t$ narrows to a special constant \top by innermost narrowing $\stackrel{i}{\rightsquigarrow}_{\mathcal{R}}$ with a substitution θ (i.e., $(s \twoheadrightarrow t) \stackrel{i}{\rightsquigarrow}_{\theta,\mathcal{R}} \top$ and thus $\theta s \stackrel{c}{\rightarrow}_{\mathcal{R}} \theta t$). Note that \twoheadrightarrow is considered a binary symbol, $(x \twoheadrightarrow x) \to \top$ is assumed to be implicitly included in \mathcal{R} , and $\stackrel{c}{\rightarrow}_{\mathcal{R}}$ denotes the constructor-based rewriting step which applies rewrite rules to basic terms. Such a narrowing tree can be the enumeration of substitutions obtained by innermost narrowing of \mathcal{R} to \top . The idea of narrowing trees has been extended to finite representations of SLD trees for logic programs [21].

Using narrowing trees, it is easy to see that there is no substitution θ such that $x < x \stackrel{i}{\rightsquigarrow} _{\theta,\mathcal{R}_1}^*$ true, and hence the above four critical pairs with $x < x \rightarrow$ true are infeasible. Let us now consider to prove infeasibility of $x < y \rightarrow$ true, $y < x \rightarrow$ true. A narrowing tree for $x < y \rightarrow$ true & $y < x \rightarrow$ true can be

http://cops.uibk.ac.at/results/?y=2016&c=CTRS

² http://cops.uibk.ac.at/results/?y=2017-full-run&c=CTRS

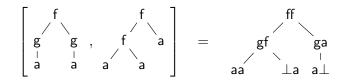


Figure 1: the coding of f(g(a), g(a)) and f(f(a, a), a).

represented by the following grammar representation [20, 19] that can be considered an RTG (Section 4):

We denote by \mathcal{G}_1 the RTG with the initial non-terminal $\Gamma_{x < y \rightarrow \text{true} \& y < x \rightarrow \text{true}}$, the other non-terminals $\Gamma_{x < y \rightarrow \text{true}}, \Gamma_{y < x \rightarrow \text{true}}$, and the above production rules. We also denote by \mathcal{P}_1 the set of the above production rules, i.e., (1). Substitutions are considered constants, and the RTG generates terms over $\&, \emptyset$, •, REC, and substitutions. The binary symbols • and & are interpreted by standard composition and *parallel composition* [11, 23], respectively. Parallel composition \Uparrow of two substitutions returns a most general unifier of the substitutions if the substitutions are unifiable (see Definition 4.2). For example, $\{y' \mapsto a, y \mapsto a\} \Uparrow \{y' \mapsto y\}$ returns $\{y' \mapsto a, y \mapsto a\}$ and $\{y' \mapsto a, y \mapsto b\} \Uparrow \{y' \mapsto y\}$ fails. The symbol REC is used for recursion, which is interpreted as standard composition of a renaming and a substitution recursively generated.

In our previous work [19], to show the emptiness of the set of substitutions generated from e.g., $\Gamma_{x < y \rightarrow \text{true}} \& \Gamma_{y < x \rightarrow \text{true}}$, we transform the grammar representation to an RTG that generates the ranges of substitutions w.r.t. a single variable. For example, for *x*, the production rules (1) of \mathcal{G}_1 is transformed into the following production rules:

$$\Gamma^{x}_{x < y \twoheadrightarrow \mathsf{true}} \to 0 \mid \mathsf{s}(\Gamma^{x}_{x < y \twoheadrightarrow \mathsf{true}}) \qquad \Gamma^{x}_{y < x \twoheadrightarrow \mathsf{true}} \to \Gamma^{x}_{x < y \twoheadrightarrow \mathsf{true}}$$

Since we focus on *x* only, both $\Gamma_{x < y \rightarrow \text{true}}^x$ and $\Gamma_{y < x \rightarrow \text{true}}^x$ generate $\{s^n(0) \mid n \ge 0\}$, and we cannot prove that there is no substitution generated from $\Gamma_{x < y \rightarrow \text{true}} \& \Gamma_{y < x \rightarrow \text{true}}$.

In this paper, we aim at showing that there is no substitution generated by (1) from $\Gamma_{x < y \rightarrow \text{true}\&y < x \rightarrow \text{true}}$, i.e., showing that $L(\mathcal{G}_1, \Gamma_{x < y \rightarrow \text{true}}) \cap L(\mathcal{G}_1, \Gamma_{y < x \rightarrow \text{true}}) = \emptyset$. To this end, under a certain syntactic condition, we show a transformation of the grammar representation of a narrowing tree into an RTG that generates the ranges of substitutions generated by the grammar representation (Section 5). More precisely, using the idea of *coding* for tuples of ground terms [2, Section 3.2.1] (see Figure 1), we extend a transformation in [19] w.r.t. a single variable to two variables (Section 3). It is straightforward to further extend the transformation to three or more variables. We do not explain how to, given a constructor 3-SDCTRS, construct (the grammar representation of) a narrowing trees, and concentrate on how to transform a grammar representation into an RTG that generates the domain of substitutions generated by the grammar representation.

One may think that tuples of terms are enough for our goal. However, substitutions are generated by standard compositions, and tuples makes us introduce composition of tuples. For example, the range of $\sigma = \{x \mapsto f(x',g(a)), y \mapsto f(y',a)\}$ is represented as a tuple tup₂(f(x',g(a)), f(y',a)), where tup₂ is a binary symbol for tuples of two terms. To apply $\theta = \{x' \mapsto g(a), y' \mapsto f(a,a)\}$ to the tuple, we reconstruct a tuple from tup₂(f(x',g(a)),f(y',a)) and θ . On the other hand, the coding of terms makes us avoid the reconstruction and use standard composition of substitutions to compute the ranges of composed substitutions— σ and θ can be represented by $\{xy \mapsto ff(x'y',ga(a\perp))\}$ and $\{x'y' \mapsto gf(aa, \perp a)\}$, respectively, where xy,x'y' are considered single variables.

2 Preliminaries

In this section, we recall basic notions and notations of term rewriting [1, 22] and regular tree grammars [2].

Throughout the paper, we use \mathcal{V} as a countably infinite set of *variables*. Let \mathcal{F} be a *signature*, a finite set of *function symbols* f each of which has its own fixed arity, denoted by arity(f). We often write $f/n \in \mathcal{F}$ instead of "an *n*-ary symbol $f \in \mathcal{F}$ ", and so on. The set of *terms* over \mathcal{F} and $V (\subseteq \mathcal{V})$ is denoted by $\mathcal{T}(\mathcal{F}, \mathcal{V})$, and $\mathcal{T}(\mathcal{F}, \emptyset)$, the set of *ground* terms, is abbreviated to $\mathcal{T}(\mathcal{F})$. The set of variables appearing in any of terms t_1, \ldots, t_n is denoted by $\mathcal{V}ar(t_1, \ldots, t_n)$. We denote the set of positions of a term t by $\mathcal{P}os(t)$. For a term t and a position p of t, the *subterm* of t at p is denoted by $t|_p$. The function symbol at the *root* position ε of a term t is denoted by root(t). Given terms s, t and a position p of s, we denote by $s[t]_p$ the term obtained from s by replacing the subterm $s|_p$ at p by t.

A substitution σ is a mapping from variables to terms such that the number of variables x with $\sigma(x) \neq x$ is finite, and is naturally extended over terms. The *domain* and *range* of σ are denoted by $\mathcal{D}om(\sigma)$ and $\mathcal{R}an(\sigma)$, respectively. The set of variables in $\mathcal{R}an(\sigma)$ is denoted by $\mathcal{VR}an(\sigma)$. We may denote σ by $\{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$ if $\mathcal{D}om(\sigma) = \{x_1, \ldots, x_n\}$ and $\sigma(x_i) = t_i$ for all $1 \leq i \leq n$. The *identity* substitution is denoted by *id*. The set of substitutions that range over a signature \mathcal{F} and a set V of variables is denoted by $Subst(\mathcal{F}, V)$: $Subst(\mathcal{F}, V) = \{\sigma \mid \sigma \text{ is a substitution}, \mathcal{R}an(\sigma) \subseteq \mathcal{T}(\mathcal{F}, V)\}$. The application of a substitution σ to a term t is abbreviated to σt , and σt is called an *instance* of t. Given a set V of variables, $\sigma|_V$ denotes the *restricted* substitution of σ w.r.t. V: $\sigma|_V = \{x \mapsto \sigma x \mid x \in \mathcal{D}om(\sigma) \cap V\}$. A substitution σ is called a *renaming* if σ is a bijection on \mathcal{V} . The *composition* $\theta \cdot \sigma$ (simply $\theta \sigma$) of substitutions σ and θ is defined as $(\theta \cdot \sigma)(x) = \theta(\sigma(x))$. A substitution σ is called *idempotent* if $\sigma\sigma = \sigma$ (i.e., $\mathcal{D}om(\sigma) \cap \mathcal{VR}an(\sigma) = \emptyset$). A substitution σ is called *more general than* a substitution θ , written by $\sigma \leq \theta$, if there exists a substitution δ such that $\delta\sigma = \theta$. A finite set E of term equations $s \approx t$ is called *unifiable* if there exists a *unifier* of E such that $\sigma s = \sigma t$ for all term equation $s \approx t$ in E. A most general unifieble. The application of a substitution θ to E is defined as $\theta(E) = \{\theta s \approx \theta t \mid s \approx t \in E\}$.

An oriented conditional rewrite rule over a signature \mathcal{F} is a triple (ℓ, r, c) , denoted by $\ell \to r \leftarrow c$, such that the *left-hand side* ℓ is a non-variable term in $\mathcal{T}(\mathcal{F}, \mathcal{V})$, the *right-hand side* r is a term in $\mathcal{T}(\mathcal{F}, \mathcal{V})$, and the *conditional part* c is a sequence $s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k$ of term pairs $(k \ge 0)$ where $s_1, t_1, \ldots, s_k, t_k \in \mathcal{T}(\mathcal{F}, \mathcal{V})$. In particular, a conditional rewrite rule is called *unconditional* if the conditional part is the empty sequence (i.e., k = 0), and we may abbreviate it to $\ell \to r$. Variables in $\mathcal{V}ar(r, c) \setminus \mathcal{V}ar(\ell)$ are called *extra variables* of the rule. An *oriented conditional term rewriting system* (a CTRS, for short) over \mathcal{F} is a set of oriented conditional rewrite rules over \mathcal{F} . A CTRS is called an (unconditional) *term rewriting system* (TRS) if every rule $\ell \to r \leftarrow c$ in the CTRS is unconditional and satisfies $\mathcal{V}ar(\ell) \supseteq \mathcal{V}ar(r)$. The *reduction relation* $\to_{\mathcal{R}}$ of a CTRS \mathcal{R} is defined as $\to_{\mathcal{R}} = \bigcup_{n\geq 0} \to_{(n),\mathcal{R}}$, where $\to_{(0),\mathcal{R}} = \emptyset$, and $\to_{(i+1),\mathcal{R}} = \{(s[\sigma\ell]_p, s[\sigma r]_p) \mid s \in \mathcal{T}(\mathcal{F}, \mathcal{V}), \ \ell \to r \leftarrow s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k \in \mathcal{R}, \ \sigma s_1 \to_{(i),\mathcal{R}}^* \ \sigma t_1, \ldots, \ \sigma s_k \to_{(i),\mathcal{R}}^* \ \sigma t_k\}$ for $i \geq 0$. To specify the position where the rule is applied, we may write $\to_{p,\mathcal{R}}$ instead of $\to_{\mathcal{R}}$. The *underlying unconditional system* $\{\ell \to r \mid \ell \to r \leftarrow c \in \mathcal{R}\}$ of \mathcal{R} is denoted by \mathcal{R}_u . A term *t* is called a *normal form* (of \mathcal{R}) if *t* is irreducible w.r.t. \mathcal{R} . A substitution σ is called *normalized* (w.r.t. \mathcal{R}) if σx is a normal form of \mathcal{R} for each variable $x \in \mathcal{D}om(\sigma)$. A CTRS \mathcal{R} is called *Type 3* (3-CTRS, for short) if every rule $\ell \to r \Leftarrow c \in \mathcal{R}$ satisfies that $\mathcal{V}ar(r) \subseteq \mathcal{V}ar(\ell, c)$. $\mathcal{V}ar(s_i) \subseteq \mathcal{V}ar(\ell, t_1, \dots, t_{i-1})$ for all $1 \le i \le k$.

The sets of *defined symbols* and *constructors* of a CTRS \mathcal{R} over a signature \mathcal{F} are denoted by $\mathcal{D}_{\mathcal{R}}$ and $\mathcal{C}_{\mathcal{R}}$, respectively: $\mathcal{D}_{\mathcal{R}} = \{root(\ell) \mid \ell \rightarrow r \leftarrow c \in \mathcal{R}\}$ and $\mathcal{C}_{\mathcal{R}} = \mathcal{F} \setminus \mathcal{D}_{\mathcal{R}}$. Terms in $\mathcal{T}(\mathcal{C}_{\mathcal{R}}, \mathcal{V})$ are called *constructor terms of* \mathcal{R} . A substitution in $Subst(\mathcal{C}_{\mathcal{R}}, \mathcal{V})$ is called a *constructor substitution of* \mathcal{R} . A term of the form $f(t_1, \ldots, t_n)$ with $f/n \in \mathcal{D}_{\mathcal{R}}$ and $t_1, \ldots, t_n \in \mathcal{T}(\mathcal{C}_{\mathcal{R}}, \mathcal{V})$ is called *basic*. A CTRS \mathcal{R} is called a *constructor system* if for every rule $\ell \rightarrow r \leftarrow c$ in \mathcal{R} , ℓ is basic. A 3-DCTRS \mathcal{R} is called *syntactically deterministic* (an SDCTRS, for short) if for every rule $\ell \rightarrow r \leftarrow s_1 \twoheadrightarrow t_1, \ldots, s_k \twoheadrightarrow t_k \in \mathcal{R}$, every t_i is a constructor term or a ground normal form of \mathcal{R}_u .

A CTRS \mathcal{R} is called *operationally terminating* if there are no infinite well-formed trees in a certain logical inference system [15]—operational termination means that the evaluation of conditions must either successfully terminate or fail in finite time. Two terms *s* and *t* are said to be *joinable*, written as $s \downarrow_{\mathcal{R}} t$, if there exists a term *u* such that $s \to_{\mathcal{R}}^* u \leftarrow_{\mathcal{R}}^* t$. A CTRS \mathcal{R} is called *confluent* if $t_1 \downarrow_{\mathcal{R}} t_2$ for any terms t_1, t_2 with $t_1 \leftarrow_{\mathcal{R}}^* \cdot \to_{\mathcal{R}}^* t_2$.

We denote a pair of terms s,t by $s \rightarrow t$ (not an equation $s \approx t$) because we analyze conditions of rewrite rules and distinguish the left- and right-hand sides of a pair $s \rightarrow t$. In addition, we deal with pairs of terms as terms by considering \rightarrow a binary function symbol. For this reason, we apply many notions for terms to pairs of terms without notice. For readability, when we deal with $s \rightarrow t$ as a term, we often bracket it: $(s \rightarrow t)$. As in [17], we assume that any CTRS in this paper implicitly includes the rule $(x \rightarrow x) \rightarrow \top$ where \top is a special constant. The rule $(x \rightarrow x) \rightarrow \top$ is used to test equivalence between two terms t_1, t_2 via $t_1 \rightarrow t_2$.

To deal with a conjunction of pairs e_1, \ldots, e_k of terms $(e_i \text{ is either } s_i \rightarrow t_i \text{ or } \top)$ as a term, we write $e_1 \& \cdots \& e_k$ by using an associative binary symbol &. We call such a term an *equational term*. Unlike [20], to avoid & to be a defined symbol, we do not use any rule for &, e.g., $(\top \& x) \rightarrow x$. Instead of derivations ending with \top , we consider derivations that end with terms in $\mathcal{T}(\{\top,\&\})$. We assume that none of &, \rightarrow , or \top is included in the range of any substitution below.

In the following, for a constructor 3-SDCTRS \mathcal{R} , a pair $s \rightarrow t$ of terms is called a *goal* of \mathcal{R} if the left-hand side *s* is either a constructor term or a basic term and the right-hand side *t* is a constructor term. An equational term is called a *goal clause* of \mathcal{R} if it is a conjunction of goals for \mathcal{R} . Note that for a goal clause *T*, any instance θT with θ a constructor substitution is a goal clause.

Example 2.1 The equational term $x < y \rightarrow$ true & $y < x \rightarrow$ true is a goal clause of \mathcal{R}_1 .

A regular tree grammar (an RTG, for short) is a quadruple $\mathcal{G} = (S, \mathcal{N}, \mathcal{F}, \mathcal{P})$ such that \mathcal{F} is a signature, \mathcal{N} is a finite set of *non-terminals* (constants not in \mathcal{F}), $S \in \mathcal{N}$, and \mathcal{P} is a finite set of *production rules* of the form $A \to \beta$ with $A \in \mathcal{N}$ and $\beta \in \mathcal{T}(\mathcal{F} \cup \mathcal{N})$. Given a non-terminal $S' \in \mathcal{N}$, the set $\{t \in \mathcal{T}(\mathcal{F}) \mid S' \to_{\mathcal{P}}^* t\}$ is the *language generated by* \mathcal{G} from S', denoted by $L(\mathcal{G}, S')$. The *initial* non-terminal S does not play an important role in this paper. A *regular tree language* is a language generated by an RTG from one of its non-terminals. The class of regular tree languages is equivalent to the class of *recognizable tree languages* which are recognized by *tree automata*. This means that the *intersection* (*non-)emptiness problem* for regular tree languages is decidable.

Example 2.2 The RTG $\mathcal{G}_2 = (X, \{X, X'\}, \{0/0, s/1\}, \{X \to 0, X \to s(X'), X' \to s(X)\})$ generates the sets of even and odd numbers over 0 and s from X and X', respectively: $L(\mathcal{G}_2, X) = \{s^{2n}(0) \mid n \ge 0\}$ $(=L(\mathcal{G}_2))$ and $L(\mathcal{G}_2, X') = \{s^{2n+1}(0) \mid n \ge 0\}$.

3 Coding of Tuples of Ground Terms

In this section, we introduce the notion of *coding* of tuples of ground terms [2, Section 3.2.1]. To simplify discussions, we consider pairs of terms.

Let \mathcal{F} be a signature. We prepare the signature $\mathcal{F}' = (\mathcal{F} \cup \{\bot\})^2$, where \bot is a new constant. For symbols $f_1, f_2 \in \mathcal{F}$, we denote the function symbol $(f_1, f_2) \in \mathcal{F}'$ by f_1f_2 , and the arity of f_1f_2 is $\max(arity(f_1), arity(f_2))$. The coding of the pair of ground terms t_1, t_2 is a ground term, denoted by $[t_1, t_2]$, over \mathcal{F}' . As described in [2, Section 3.2.1], the basic idea of coding is to stack function symbols as illustrated in Figure 1. The coding of pairs of ground terms, $[\cdot, \cdot]$, is recursively defined as follows:

- $[f(s_1,...,s_m), g(t_1,...,t_n)] = fg([s_1, t_1],...,[s_m, t_m], [\bot, t_{m+1}],..., [\bot, t_n])$ if $m \le n$, and
- $[f(s_1,...,s_m), g(t_1,...,t_n)] = fg([s_1, t_1],...,[s_n, t_n], [s_{n+1}, \bot],...,[s_m, \bot])$ if m > n.

Note that $\mathcal{P}os([t_1, t_2]) = \mathcal{P}os(t_1) \cup \mathcal{P}os(t_2)$, and for i = 1, 2 and $p \in \mathcal{P}os([t_1, t_2])$, if $p \notin \mathcal{P}os(t_i)$, then \perp is complemented for t_i .

Example 3.1 As in Figure 1, $[f(g(a), g(a)), f(f(a, a), a)] = ff(gf(aa, \perp a), ga(a\perp)).$

4 Grammar Representations for Sets of Idempotent Substitutions

In this section, we briefly introduce grammar representations that define sets of idempotent substitutions. We follow the formalization in [19], which is based on *success set equations* in [20]. Since substitutions derived by narrowing steps are assumed to be idempotent, we only deal with idempotent substitutions which introduce only *fresh* variables which do not appear in any previous term.

In the following, a renaming ξ is used to rename a particular term *t* and we assume that $\xi|_{Var(t)}$ is injective on Var(t). For this reason, we write $\xi|_{Var(t)}$ instead of ξ , and call $\xi|_{Var(t)}$ a renaming for *t* (simply, a renaming).

We first introduce terms to represent idempotent substitutions computed using \cdot and \uparrow . We prepare the signature Σ consisting of the following symbols:

• idempotent substitutions,	(basic elements)
• a constant \emptyset ,	(the empty set/non-existence)
• an associative binary symbol •,	(standard composition)
• an associative binary symbol &, and	(parallel composition)
• a binary symbol REC.	(recursion with renaming)

We use infix notation for \bullet and &, and may omit brackets with the precedence such that \bullet has a higher priority than &.

We deal with terms over Σ and some constants used for non-terminals of grammar representations, where we allow such constants to only appear in the first argument of REC. Note that a term without any constant may appear in the first argument of REC. Given a finite set \mathcal{N} of constants, we denote the set of such terms by $\mathcal{T}(\Sigma \cup \mathcal{N})$. We assume that each constant in \mathcal{N} has a term t (possibly a goal clause) as subscript such as Γ_t . For an expression REC(Γ_t , δ), the role of Γ_t is recursion to generate terms in $\mathcal{T}(\Sigma)$. To reuse substitutions generated by recursion, we connect them with other substitutions via some renaming δ . For this reason, we restrict the second argument of REC to renamings and we require each term REC(Γ_t , δ) to satisfy $\mathcal{VRan}(\delta) = \mathcal{Var}(t)$. *Example 4.1* The following are instances of terms in $\mathcal{T}(\Sigma)$:

- $\{y \mapsto 0\} \bullet \{x \mapsto \mathsf{s}(y)\},\$
- $(\{x' \mapsto \mathsf{s}(y)\} \bullet \{x \mapsto x'\}) \& \{x \mapsto \mathsf{s}(\mathsf{s}(z))\},\$
- $(\emptyset \& \{y \mapsto z\}) \bullet \{x \mapsto \mathsf{s}(y)\}$, and
- REC({ $x \mapsto 0, y \mapsto s(y')$ }, { $x' \mapsto x, y' \mapsto y$ }) { $y \mapsto s(x')$ }.

Note that substitutions $\{y \mapsto 0\}$, $\{x \mapsto s(y)\}$, $\{x' \mapsto s(y)\}$, $\{x \mapsto x'\}$, $\{x \mapsto s(s(z))\}$, $\{y \mapsto z\}$, $\{x \mapsto 0, y \mapsto s(y')\}$, $\{x' \mapsto x, y' \mapsto y\}$, $\{y \mapsto s(x')\}$ are considered constants.

Next, we recall *parallel composition* \Uparrow of idempotent substitutions [11, 23], which is one of the most important key operations to enable us to construct *finite* narrowing trees. Given a substitution $\theta = \{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$, we denote the set of term equations $\{x_1 \approx t_1, \ldots, x_n \approx t_n\}$ by $\hat{\theta}$.

Definition 4.2 (parallel composition \Uparrow **[23])** *Let* θ_1 *and* θ_2 *be idempotent substitutions. Then, we define* \Uparrow *as follows:* $\theta_1 \Uparrow \theta_2 = mgu(\hat{\theta}_1 \cup \hat{\theta}_2)$ *if* $\hat{\theta}_1 \And \hat{\theta}_2$ *is unifiable, and otherwise,* $\theta_1 \Uparrow \theta_2 = fail$ *. Note that we define* $\theta_1 \Uparrow \theta_2 = fail$ *if* θ_1 *or* θ_2 *is not idempotent. Parallel composition is extended to sets* Θ_1, Θ_2 *of idempotent substitutions in the natural way:* $\Theta_1 \Uparrow \Theta_2 = \{\theta_1 \Uparrow \theta_2 \mid \theta_1 \in \Theta_1, \theta_2 \in \Theta_2, \theta_1 \Uparrow \theta_2 \neq fail\}.$

We often have two or more substitutions that can be results of $\theta_1 \Uparrow \theta_2$ ($\neq fail$), while most general unifiers are unique up to variable renaming. In this paper, to simplify the semantics of grammar representations for substitutions, we adopt an idempotent substitution σ with $\mathcal{D}om(\theta_1) \cup \mathcal{D}om(\theta_2) \subseteq \mathcal{D}om(\sigma)$ as a result of $\theta_1 \Uparrow \theta_2$ ($\neq fail$). Note that most general unifiers we can adopt as results of $\theta_1 \Uparrow \theta_2$ under the convention are still not unique, while they are unique up to variable renaming.

Example 4.3 The parallel composition $\{x \mapsto s(z), y \mapsto z\} \Uparrow \{x \mapsto w\}$ may return $\{x \mapsto s(z), y \mapsto z, w \mapsto s(z)\}$, but we do not allow $\{x \mapsto s(y), z \mapsto y, w \mapsto s(y)\}$ as a result. On the other hand, $\{x \mapsto s(z), y \mapsto z\} \Uparrow \{x \mapsto y\}$ fails.

When we compute $\sigma_1 \Uparrow \sigma_2$, we assume that $\mathcal{VRan}(\sigma_1) \cap \mathcal{VRan}(\sigma_2) = \emptyset$. To satisfy this assumption explicitly in the semantics for $\mathcal{T}(\Sigma)$, we introduce an operation $fresh_{\delta}(\cdot)$ of substitutions to make a substitution introduce only variables that do not appear in $\mathcal{Dom}(\delta) \cup \mathcal{VRan}(\delta)$: for substitutions σ, δ , we define $fresh_{\delta}(\sigma)$ by $(\xi \cdot \sigma)|_{\mathcal{Dom}(\sigma)}$ where ξ is a renaming such that $\mathcal{Dom}(\xi) \supseteq \mathcal{VRan}(\sigma)$ and $\mathcal{VRan}(\xi|_{\mathcal{VRan}(\sigma)}) \cap (\mathcal{Dom}(\delta) \cup \mathcal{VRan}(\delta) \cup \mathcal{Dom}(\sigma)) = \emptyset$. The subscript δ of $fresh_{\delta}(\cdot)$ is used to specify freshness of variables—we say that a variable x is *fresh* w.r.t. a set X of variables if $x \notin X$.

A term *e* in $\mathcal{T}(\Sigma)$ defines a substitution. The semantics of terms in $\mathcal{T}(\Sigma)$ is inductively defined as follows:

- $\llbracket \theta \rrbracket = \theta$ if θ is a substitution,
- $\llbracket e_1 \bullet e_2 \rrbracket = \llbracket e_1 \rrbracket \cdot \llbracket e_2 \rrbracket$ if $\llbracket e_2 \rrbracket \neq fail$ and $\llbracket e_1 \rrbracket \neq fail$,
- $\llbracket e_1 \& e_2 \rrbracket = (\theta_1 \Uparrow \theta_2)|_{\mathcal{D}om(\theta_1) \cup \mathcal{D}om(\theta_2)}$ if $\llbracket e_1 \rrbracket \neq fail$ and $\llbracket e_2 \rrbracket \neq fail$, where $\theta_1 = \llbracket e_1 \rrbracket$ and $\theta_2 = fresh_{\theta_1}(\llbracket e_2 \rrbracket)$,
- $[\operatorname{REC}(e, \delta)] = (\operatorname{fresh}_{\delta}([e]) \cdot \delta)|_{\mathcal{D}om(\delta)}$ if $[e] \neq fail$ and $\mathcal{VRan}(\delta) = \mathcal{D}om([e])$, and
- otherwise, $\llbracket e \rrbracket = fail$ (e.g., $\llbracket \varnothing \rrbracket = fail$).

Notice that Γ_t is not included in $\mathcal{T}(\Sigma)$, and thus, $[\![\Gamma_t]\!]$ is not defined. Since \Uparrow may fail, we allow to have *fail*, e.g., $[\![\{y \mapsto s(x)\}\!] \bullet \{x \mapsto y\}\!\& \{z \mapsto 0\}\!]] = fail$. The number of variables appearing in an RTG defined below is finite. However, we would like to use RTGs to define infinitely many substitutions such that the maximum number of variables we need cannot be fixed. To solve this problem, in the definition of $[\![REC(e, \delta)]\!]$, we introduced the operation $fresh_{\delta}(\cdot)$ that make all variables introduced by $[\![e]\!]$ *fresh* w.r.t. $Dom(\delta) \cup \mathcal{VRan}(\delta)$. In [20], this operation is implicitly considered, but in [19], REC is explicitly introduced to the syntax in order to convert terms in $\mathcal{T}(\Sigma)$ precisely. To assume $\mathcal{VRan}([\![e_1]\!]) \cap \mathcal{VRan}([\![e_2]\!]) = \emptyset$ for $[\![e_1 \& e_2]\!]$, we also introduced *fresh*_{\theta_1}(\cdot) in the case of $[\![e_1 \& e_2]\!]$.

Example 4.4 The expressions in Example 4.1 are interpreted as follows:

- $\llbracket \{y \mapsto 0\} \bullet \{x \mapsto \mathsf{s}(y)\} \rrbracket = \{x \mapsto \mathsf{s}(0), y \mapsto 0\},\$
- $\llbracket (\{x' \mapsto \mathsf{s}(y)\} \bullet \{x \mapsto x'\}) \& \{x \mapsto \mathsf{s}(\mathsf{s}(z))\} \rrbracket = \{x \mapsto \mathsf{s}(\mathsf{s}(z)), x' \mapsto \mathsf{s}(\mathsf{s}(z))\},\$
- $\llbracket (\varnothing \& \{y \mapsto z\}) \bullet \{x \mapsto \mathsf{s}(y)\} \rrbracket = fail, \text{ and }$
- $[\![\operatorname{REC}(\{x \mapsto 0, \ y \mapsto \mathsf{s}(y')\}, \{x' \mapsto x, \ y' \mapsto y\}) \bullet \{y \mapsto \mathsf{s}(x')\}]\!] = \{x' \mapsto 0, \ y' \mapsto \mathsf{s}(y''), \ y \mapsto \mathsf{s}(0)\}.$

To define sets of idempotent substitutions, we adopt RTGs. In the following, we drop the third component from grammars constructed below because the third one is fixed to Σ with a finite number of substitutions that are clear from production rules. A *substitution-set grammar* (SSG) for a term t_0 is an RTG $\mathcal{G} = (\Gamma_{t_0}, \mathcal{N}, \mathcal{P})$ such that \mathcal{N} is a finite set of non-terminals $\Gamma_t, \Gamma_{t_0} \in \mathcal{N}$, and \mathcal{P} is a finite set of production rules of the form $\Gamma_t \to \beta$ with $\beta \in \mathcal{T}(\Sigma \cup \mathcal{N})$. Note that $L(\mathcal{G}, \Gamma_t) = \{e \in \mathcal{T}(\Sigma) \mid \Gamma_t \to_{\mathcal{G}}^* e\}$ for each $\Gamma_t \in \mathcal{N}$, and the numbers of variables appearing in $L(\mathcal{G}, \Gamma_t)$ is finite. The set of substitutions generated by \mathcal{G} from $\Gamma_t \in \mathcal{N}$ is defined as $[[L(\mathcal{G}, \Gamma_t)]] = \{[[e]] \mid e \in L(\mathcal{G}, \Gamma_t), [[e]] \neq fail\}$. Note that the number of variables in $\bigcup_{\theta \in [[L(\mathcal{G}, \Gamma_t)]]} \mathcal{VR}an(\theta)$ may be infinite because of the interpretation for REC.

Example 4.5 The SSG $\mathcal{G}_3 = (\Gamma_x, \{\Gamma_x, \Gamma_y\}, \{\Gamma_x \to \{x \mapsto 0\} \mid \text{REC}(\Gamma_y, \{x' \mapsto y\}) \bullet \{x \mapsto \mathsf{s}(x')\}, \Gamma_y \to \text{REC}(\Gamma_x, \{x' \mapsto x\}) \bullet \{y \mapsto \mathsf{s}(x')\}\}$ generates a set of expressions to define substitutions replacing x, y by even numbers over 0/0 and $\mathsf{s}/1$. We have that $L(\mathcal{G}_3) = L(\mathcal{G}_3, \Gamma_x) = \{\{x \mapsto 0\}, \text{REC}((\text{REC}(\{x \mapsto 0\}, \{x' \mapsto x\}) \bullet \{y \mapsto \mathsf{s}(x')\}), \{x' \mapsto y\}) \bullet \{x \mapsto \mathsf{s}(x')\}, \ldots\}$, and $[[L(\mathcal{G}_3, \Gamma_x)]] = \{\{x \mapsto \mathsf{s}^{2n}(0)\} \mid n \ge 0\}$.

5 Transforming SSGs into RTGs Generating Ranges of Substitutions

In this section, given a goal clause *T* and two variables x_1, x_2 in *T*, we show a transformation of an SSG $\mathcal{G} = (\Gamma_{T_0}, \mathcal{N}, \mathcal{P})$ into an RTG \mathcal{G}' such that $L(\mathcal{G}', \Gamma_T^{x_1, x_2}) \supseteq \{ [\xi \theta x_1, \xi \theta x_2] \mid \theta \in L(\mathcal{G}, \Gamma_T), \xi \in Subst(\mathcal{C}_{\mathcal{R}}), \ \mathcal{V}ar(\theta x_1, \theta x_2) \subseteq \mathcal{D}om(\xi) \}$. Note that *T* does not have to be T_0 . The transformation is applicable to SSGs satisfying some certain syntactic condition shown later. In the following, we aim at showing that $L(\mathcal{G}_1, \Gamma_{x < y \rightarrow \text{true}}) \cap L(\mathcal{G}_1, \Gamma_{y < x \rightarrow \text{true}}) = \emptyset$.

Let \mathcal{G} be an SSG $(\Gamma_{T_0}, \mathcal{N}, \mathcal{P})$ and T a goal clause such that $\Gamma_T \in \mathcal{N}$. We denote by $\mathcal{P}|_{\Gamma_T}$ the set of production rules that are reachable from Γ_T . We assume that any rule in $\mathcal{P}|_{\Gamma_T}$ is of the following form:

$$\Gamma_{T'} \rightarrow \theta_1 \mid \cdots \mid \theta_m \mid \operatorname{REC}(\Gamma_{T_1}, \delta_1) \bullet \theta_{m+1} \mid \cdots \mid \operatorname{REC}(\Gamma_{T_n}, \delta_n) \bullet \theta_{m+n}$$

where $\theta_1, \ldots, \theta_{m+n}$ are idempotent substitutions. Note that $\Gamma_{T'} \to \text{REC}(\Gamma_{T''}, \delta)$ is considered $\Gamma_{T'} \to \text{REC}(\Gamma_{T''}, \delta) \bullet id$. In addition, we assume that for each $\Gamma_{T'} \to \text{REC}(\Gamma_{T_i}, \delta_i) \bullet \theta_{m+i}$ with $1 \le i \le n$, for all variables x, y in T', for all positions $p \in \mathcal{P}os(\delta \theta_{m+i}x) \cap \mathcal{P}os(\delta \theta_{m+i}y)$, all of the following hold:

• if $(\delta \theta_{m+i}x)|_p \in \mathcal{V}ar(T_i)$, then $(\delta \theta_{m+i}y)|_p \in \mathcal{V}ar(T_i) \cup \mathcal{T}(\mathcal{C}_{\mathcal{R}}, \mathcal{V} \setminus \mathcal{V}ar(T_i))$, and

• if $(\delta \theta_{m+i}y)|_p \in \mathcal{V}ar(T_i)$, then $(\delta \theta_{m+i}x)|_p \in \mathcal{V}ar(T_i) \cup \mathcal{T}(\mathcal{C}_{\mathcal{R}}, \mathcal{V} \setminus \mathcal{V}ar(T_i))$.

This assumption implies that for for all variables x, y in T', for all positions $p \in \mathcal{P}os(\delta \theta_{m+i}x) \cap \mathcal{P}os(\delta \theta_{m+i}y)$, the terms $(\delta \theta_{m+i}x)|_p$ and $(\delta \theta_{m+i}y)|_p$ satisfy one of the following:

- (a) both are rooted by function symbols,
- (b) both are variables in $\mathcal{V}ar(T_i)$,
- (c) one is a variable in $Var(T_i)$ and the other is a term in $T(C_R, V \setminus Var(T_i))$, or
- (d) both are terms in $\mathcal{T}(\mathcal{C}_{\mathcal{R}}, \mathcal{V} \setminus \mathcal{V}ar(T_i))$.

For example, both $\mathcal{P}_1|_{\Gamma_{x<y\to true}}$ and $\mathcal{P}_1|_{\Gamma_{y<\to true}}$ satisfy the above assumption. Our idea of the extension is the use of coding; Roughly speaking, for $\Gamma_{T'} \to \text{REC}(\Gamma_{T_i}, \delta_i) \bullet \theta_{m+i}$ with $1 \le i \le n$, for all variables x, y in T', we apply *coding* to $\delta \theta_{m+i}x$ and $\delta \theta_{m+i}y$. A variable in $\mathcal{V}ar(T_i)$, which is instantiated by substitutions generated from Γ_{T_i} , may prevent us from constructing production rules. For this reason, we expect any variable³ in $\mathcal{V}ar(\delta \theta_{m+i}x, \delta \theta_{m+i}y) \cap \mathcal{V}ar(T_i)$ to be coded with

- \perp (the case where the precondition " $p \in \mathcal{P}os(\delta \theta_{m+i}x) \cap \mathcal{P}os(\delta \theta_{m+1}y)$ " does not hold),
- another variable in $\mathcal{V}ar(\delta \theta_{m+i}x, \delta \theta_{m+i}y) \cap \mathcal{V}ar(T_i)$ (the case where (b) above holds), or
- a constructor term without any variable in $Var(T_i)$ (the case where (c) above holds).

Note that the following construction is applicable under the above assumption. We denote the set of constructor terms appearing in substitutions in \mathcal{P} by $Patterns(\mathcal{P})$, where such constructor terms are instantiated with A a non-terminal introduced during the transformation below: $Patterns(\mathcal{P}) =$ $\{\{x \mapsto A \mid x \in \mathcal{V}ar(s)\}(t) \mid \theta \text{ appears in } \mathcal{P}, s \in \mathcal{V}\mathcal{R}an(\theta), t \leq s\}$. We denote the set of variables appearing in \mathcal{N} by $Vars(\mathcal{N})$: $Vars(\mathcal{N}) = \bigcup_{\Gamma_{T}' \in \mathcal{N}} \mathcal{V}ar(T')$. The RTG obtained from \mathcal{G} and variables x_1, x_2 in T, denoted by $\mathcal{R}an(\mathcal{G}, T, x, y)$, is $(\Gamma_T^{x_1, x_2}, \mathcal{N}', \mathcal{P}'_1 \cup \mathcal{P}'_2 \cup \mathcal{P}_{AA} \cup \mathcal{P}_{A\perp} \cup \mathcal{P}_{\perp A})$ such that

- $\mathcal{N}' = \{ \Gamma_{T'}^{(t_1,t_2)}, \Gamma_{T'}^{(t_1,\perp)}, \Gamma_{T'}^{(\perp,t_2)} | \Gamma_{T'} \in \mathcal{N}, t_1, t_2 \in Patterns(\mathcal{P}) \} \cup \{ \Gamma_{T'}^{(x,y)}, \Gamma_{T'}^{(x,t)}, \Gamma_{T'}^{(t,y)} | x, y \in \{x_1, x_2\} \cup Vars(\mathcal{N}), \Gamma_{T'} \in \mathcal{N}, t \in Patterns(\mathcal{P}) \cup \{ \perp \} \},$
- $\mathcal{P}'_1 = \{ \Gamma_{T'}^{(t_1,t_2)} \to u \mid \Gamma_{T'} \to \theta \in \mathcal{P}, \ \Gamma_{T'}^{(t_1,t_2)} \in \mathcal{N}', \ u \in [\xi \theta t_1, \xi \theta t_2]_{\top} \}, \text{ where } \xi = \{x \mapsto A \mid x \in \mathcal{V}ar(\theta t_1, \theta t_2)\},$
- $\mathcal{P}'_2 = \{ \Gamma_{T'}^{(t_1,t_2)} \rightarrow u \mid \Gamma_{T'} \rightarrow \text{REC}(\Gamma_{T''}, \delta) \bullet \theta \in \mathcal{P}, \ \Gamma_{T'}^{(t_1,t_2)} \in \mathcal{N}', \ u \in [\xi \delta \theta t_1, \xi \delta \theta t_2]_{T''} \},$ where $\xi = \{x \mapsto A \mid x \in \mathcal{V}ar(\delta \theta t_1, \delta \theta t_2) \setminus \mathcal{V}ar(T'')\},$

•
$$\mathcal{P}_{AA} = \{ AA \to u \mid f/m, g/n \in \mathcal{C}_{\mathcal{R}}, u \in \left[f(A, \dots, A), g(A, \dots, A) \right]_{\top} \},$$

•
$$\mathcal{P}_{A\perp} = \{ A \perp \to u \mid f/m \in \mathcal{C}_{\mathcal{R}}, u \in \left[f(\overbrace{A, \dots, A}^{m}), \perp \right]_{\top} \}, \text{ and}$$

•
$$\mathcal{P}_{\perp A} = \{ \perp A \rightarrow u \mid g/n \in \mathcal{C}_{\mathcal{R}}, u \in \left\lfloor \perp, g(\widehat{A, \ldots, A}) \right\rfloor_{\top} \}$$

where $[\cdot, \cdot]_{T'}$, which takes a goal clause T' and two terms in $\mathcal{T}(\mathcal{F} \cup \{A\}, \mathcal{V}ar(T'))$ as input and returns a set of terms in $\mathcal{T}(\mathcal{F} \cup \mathcal{N}')$, is recursively defined as follows:

 $^{^{3}}$ This is not the case where (a) holds.

- $[x, y]_{T'} = \{ \Gamma_{T'}^{(x,y)} \}$, where $x, y \in \mathcal{V}$,
- $[x, A]_{T'} = \{ \Gamma_{T'}^{(x,A)} \}$, where $x \in \mathcal{V}$,
- $[x, \perp]_{T'} = \{ \Gamma_{T'}^{(x,\perp)} \}$, where $x \in \mathcal{V}$,
- $[A, y]_{T'} = \{ \Gamma_{T'}^{(A,y)} \}$, where $y \in \mathcal{V}$,
- $[\bot, y]_{T'} = \{ \Gamma_{T'}^{(\bot, y)} \}$, where $y \in \mathcal{V}$,
- $[A, A]_{T'} = \{AA\},\$
- $[A, \bot]_{T'} = \{ A \bot \},$
- $[\bot, A]_{T'} = \{ \bot A \},$
- $[\bot, g(t_1,...,t_n)]_{T'} = \{ \bot g(u_1,...,u_n) \mid 1 \le i \le n, u_i \in [\bot, t_i]_{T'} \},$
- $[f(s_1,...,s_m), \perp]_{T'} = \{ f \perp (u_1,...,u_m) \mid 1 \le i \le m, u_i \in [s_i, \perp]_{T'} \},\$
- $[A, g(t_1,...,t_n)]_{T'} = \{ fg(u_1,...,u_m,u_{m+1},...,u_n) \mid f/m \in C_{\mathcal{R}}, m < n, 1 \le i \le m, u_i \in [A, t_i]_{T'}, 1 \le j \le n-m, u_{m+j} \in [\bot, t_{m+j}]_{T'} \} \cup \{ fg(u_1,...,u_n,u_{n+1},...,u_m) \mid f/m \in C_{\mathcal{R}}, m \ge n, 1 \le i \le n, u_i \in [A, t_i]_{T'}, 1 \le j \le m-n, u_{n+j} \in [A, \bot]_{T'} \},$
- $[f(s_1,...,s_m), A]_{T'} = \{ fg(u_1,...,u_m,u_{m+1},...,u_n) \mid g/n \in C_{\mathcal{R}}, m < n, 1 \le i \le m, u_i \in [s_i, A]_{T'}, 1 \le j \le n-m, u_{m+j} \in [\bot, A]_{T'} \} \cup \{ fg(u_1,...,u_n,u_{n+1},...,u_m) \mid g/n \in C_{\mathcal{R}}, m \ge n, 1 \le i \le n, u_i \in [s_i, A]_{T'}, 1 \le j \le m-n, u_{n+j} \in [s_{n+j}, \bot]_{T'} \},$
- $[f(s_1,...,s_m), g(t_1,...,t_n)]_{T'} = \{ fg(u_1,...,u_m,u_{m+1},...,u_n) \mid 1 \le i \le m, u_i \in [s_i, t_i]_{T'}, 1 \le j \le n-m, u_{m+j} \in [\bot, t_{m+j}]_{T'} \}$ if m < n, and
- $[f(s_1,...,s_m), g(t_1,...,t_n)]_{T'} = \{ fg(u_1,...,u_n,u_{n+1},...,u_m) \mid 1 \le i \le n, u_i \in [s_i, t_i]_{T'}, 1 \le j \le m-n, u_{n+j} \in [s_{n+j}, \bot]_{T'} \}$ if $m \ge n$.

Note that the non-terminal *AA* generates $\{[t_1, t_2] | t_1, t_2 \in \mathcal{T}(\mathcal{C}_{\mathcal{R}})\}$, the non-terminal $A\perp$ generates $\{[t_1, \perp] | t_1 \in \mathcal{T}(\mathcal{C}_{\mathcal{R}})\}$, and the non-terminal $\perp A$ generates $\{[\perp, t_2] | t_2 \in \mathcal{T}(\mathcal{C}_{\mathcal{R}})\}$. Note also that we generate only production rules that are reachable from $\Gamma_T^{(x_1,x_2)}$, and drop non-terminals not appearing in the generated production rules from \mathcal{N}' .

Example 5.1 Consider $\mathcal{G}_1 = (\Gamma_{x < y \rightarrow \text{true} \& y < x \rightarrow \text{true}}, \{\Gamma_{x < y \rightarrow \text{true} \& y < x \rightarrow \text{true}}, \Gamma_{x < y \rightarrow \text{true}}, \Gamma_{y < x \rightarrow \text{true}}\}, \mathcal{P}_1)$ in Section 1. We have that

- *Patterns*(\mathcal{P}_1) = {0,s(A), A}, and
- $Vars({\Gamma_{x < y \rightarrow true \& y < x \rightarrow true}, \Gamma_{x < y \rightarrow true}, \Gamma_{y < x \rightarrow true}}) = {x, y}.$

Let us focus on $\Gamma_{x < y \rightarrow \text{true}}$ and x, y. Since neither $\Gamma_{x < y \rightarrow \text{true} \& y < x \rightarrow \text{true}}$ nor $\Gamma_{y < x \rightarrow \text{true}}$ is reachable from $\Gamma_{x < y \rightarrow \text{true}}$ by \mathcal{P}_1 , we do not take $\Gamma_{x < y \rightarrow \text{true} \& y < x \rightarrow \text{true}}$, $\Gamma_{y < x \rightarrow \text{true}}$, and their rules into account when we construct the RTG $\mathcal{R}an(\mathcal{G}_1, \Gamma_{x < y \rightarrow \text{true}}, x, y)$. The RTG $\mathcal{R}an(\mathcal{G}_1, \Gamma_{x < y \rightarrow \text{true}}, x, y) = (\Gamma_{x < y \rightarrow \text{true}}^{(x,y)}, \mathcal{N}', \mathcal{P}'_1 \cup \mathcal{P}'_2 \cup \mathcal{P}_{AA} \cup \mathcal{P}_{A\perp} \cup \mathcal{P}_{\perp A})$ is constructed as follows:

- $\mathcal{N}' = \{ \Gamma_{x < y \rightarrow \mathsf{true}}^{(0,0)}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(0,\perp)}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(\perp,0)}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(0,\mathsf{s}(A))}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(\perp,\mathsf{s}(A))}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(\perp,\mathsf{s}(A))}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(0,\mathsf{s}(A),\mathsf{s})}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(s,\mathsf{s}(A),\mathsf{s})}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(s,\mathsf{s}(A),\mathsf{s})}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(s,\mathsf{s}(A),\mathsf{s})}, \Gamma_{x < y \rightarrow \mathsf{true}}^{(\mathsf{s}(A),\mathsf{s})}, \Gamma_{x < \mathsf$
- $\mathcal{P}'_1 = \{\Gamma^{(x,y)}_{x < y \to \mathsf{true}} \to \mathsf{0s}(\bot A)\}$, because $\Gamma_{x < y \to \mathsf{true}} \to \{x \mapsto \mathsf{0}, \ y \mapsto \mathsf{s}(y_2)\} \in \mathcal{P}_1$ and $[\mathsf{0}, \mathsf{s}(A)]_{\top} = \{\mathsf{0s}(\bot A)\}$

- $\mathcal{P}'_2 = \{\Gamma^{(x,y)}_{x < y \to \mathsf{true}} \to \mathsf{ss}(\Gamma^{(x,y)}_{x < y \to \mathsf{true}})\}, \text{ because } \Gamma_{x < y \to \mathsf{true}} \to \mathsf{REC}(\Gamma_{x < y \to \mathsf{true}}, \{x_3 \mapsto x, y_3 \mapsto y\}) \bullet \{x \mapsto \mathsf{s}(x_3), y \mapsto \mathsf{s}(y_3)\} \in \mathcal{P}_1 \text{ and } [\mathsf{s}(x), \mathsf{s}(y)]_{x < y \to \mathsf{true}} = \{\mathsf{ss}(\Gamma^{(x,y)}_{x < y \to \mathsf{true}})\},$
- $\mathcal{P}_{AA} = \{ AA \rightarrow 00 \mid 0s(\perp A) \mid 0true \mid 0false \mid s0(A\perp) \mid ss(AA) \mid strue(A\perp) \mid sfalse(A\perp) \mid true0 \mid trues(\perp A) \mid truetrue \mid truefalse \mid false0 \mid falses(\perp A) \mid falsetrue \mid falsefalse \}$
- $\mathcal{P}_{A\perp} = \{ A \perp \rightarrow 0 \perp | s \perp (A \perp) | true \perp | false \perp \}, and$
- $\mathcal{P}_{\perp A} = \{ \perp A \rightarrow \perp 0 \mid \perp s(\perp A) \mid \perp true \mid \perp false \}.$

For $\Gamma_{y < x \rightarrow \text{true}}$ and x, y, we add $\Gamma_{y < x \rightarrow \text{true}}^{(x,y)} \rightarrow \Gamma_{x < y \rightarrow \text{true}}^{(y,x)}$ to the above production rules. Rules that are not reachable from $\Gamma_{x < y \rightarrow \text{true}}^{(x,y)}$ or $\Gamma_{y < x \rightarrow \text{true}}^{(x,y)}$ can be dropped from $\mathcal{R}an(\mathcal{G}_1, \Gamma_{x < y \rightarrow \text{true}}, x, y)$, obtaining an RTG, denoted by \mathcal{G}_4 , with the following production rules:

$$\begin{split} \Gamma^{(x,y)}_{x < y \rightarrow \mathsf{true}} &\to \mathsf{Os}(\bot A) \mid \mathsf{ss}(\Gamma^{(x,y)}_{x < y \rightarrow \mathsf{true}}) \\ \Gamma^{(y,x)}_{x < y \rightarrow \mathsf{true}} &\to \mathsf{sO}(A \bot) \mid \mathsf{ss}(\Gamma^{(y,x)}_{x < y \rightarrow \mathsf{true}}) \\ \Gamma^{(x,y)}_{y < x \rightarrow \mathsf{true}} &\to \Gamma^{(y,x)}_{x < y \rightarrow \mathsf{true}} \\ & A \bot \to \mathsf{O}\bot \mid \mathsf{s}\bot(A \bot) \mid \mathsf{true}\bot \mid \mathsf{false}\bot \\ & \bot A \rightarrow \bot \mathsf{O} \mid \bot \mathsf{s}(\bot A) \mid \bot \mathsf{true} \mid \bot \mathsf{false} \end{split}$$

It is easy to see that

- $L(\mathcal{G}_4, \Gamma_{x < y \twoheadrightarrow \mathsf{true}}^{(x,y)}) \subseteq \mathcal{T}(\{\mathsf{0s}, \mathsf{ss}, \bot \mathsf{0}, \bot \mathsf{s}, \bot \mathsf{true}, \bot \mathsf{false}\}),$
- $L(\mathcal{G}_4, \Gamma_{y < x \rightarrow \mathsf{true}}^{(x,y)}) \subseteq \mathcal{T}(\{\mathsf{s0}, \mathsf{ss}, \mathsf{0}\bot, \mathsf{s}\bot, \mathsf{true}\bot, \mathsf{false}\bot\}),$

and hence, there is no shared constant between the two sets. This means that

$$L(\mathcal{G}_4, \Gamma_{x < y \rightarrow \mathsf{true}}^{(x,y)}) \cap L(\mathcal{G}_4, \Gamma_{y < x \rightarrow \mathsf{true}}^{(x,y)}) = \emptyset$$

This implies that $L(\mathcal{G}_1, \Gamma_{x < y \rightarrow \text{true}}) \cap L(\mathcal{G}_1, \Gamma_{y < x \rightarrow \text{true}}) = \emptyset$. Note that the emptiness problem of RTGs is decidable, and hence we can decide the emptiness problem of $L(\mathcal{G}_4, \Gamma_{x < y \rightarrow \text{true}}^{(x,y)}) \cap L(\mathcal{G}_4, \Gamma_{y < x \rightarrow \text{true}}^{(x,y)})$.

The following example illustrates why we adopt the assumption.

Example 5.2 Consider the SSG $\mathcal{G}_5 = (\Gamma_{x \to y}, \{\Gamma_{x \to y}\}, \{\Gamma_{x \to y} \to \{x \mapsto 0, y \mapsto 0\} | \operatorname{REC}(\Gamma_{x \to y}, \{x' \mapsto x, y' \mapsto y\}) \bullet \{x \mapsto \mathsf{s}(x'), y \mapsto \mathsf{s}(\mathsf{s}(y'))\} \})$, which does not satisfy the assumption. The domains of substitutions generated by \mathcal{G}_5 w.r.t. (x, y) is $\{(\mathsf{s}^n(0), \mathsf{s}^{2n}(0)) | n \ge 0\}$ which is not recognizable. This implies that there is no RTG generating this set.

6 Conclusion

In this paper, under a certain syntactic condition, we showed a transformation of the grammar representation of a narrowing tree into an RTG that generates the ranges of substitutions generated by the grammar representation. We showed that the transformation is useful to prove infeasibility of conditional critical pairs of \mathcal{R}_1 , i.e., confluence of \mathcal{R}_1 . We showed a precise definition of the transformation, but have not proved the correctness yet. We will prove the correctness, and make an experiment to evaluate the usefulness of the transformation in e.g., proving confluence of CTRSs.

The syntactic assumption in Section 5 is a sufficient condition to, given an SSG, obtain an RTG that generates the ranges of substitutions generated by the SSG. It is not known yet whether the assumption is a necessary condition or not. We will try to clarify this point.

References

- [1] Franz Baader & Tobias Nipkow (1998): *Term Rewriting and All That.* Cambridge University Press, doi:10.1017/CBO9781139172752.
- [2] Hubert Comon, Max Dauchet, Rémi Gilleron, Florent Jacquemard, Denis Lugiez, Christof Löding, Sophie Tison & Marc Tommasi (2007): *Tree Automata Techniques and Applications*. Available on: http://www. grappa.univ-lille3.fr/tata. Release October, 12th 2007.
- [3] Nachum Dershowitz, Mitsuhiro Okada & G. Sivakumar (1988): Canonical Conditional Rewrite Systems. In Ewing L. Lusk & Ross A. Overbeek, editors: Proceedings of the 9th International Conference on Automated Deduction, Lecture Notes in Computer Science 310, Springer, pp. 538–549, doi:10.1007/BFb0012855.
- [4] Francisco Durán, Salvador Lucas, José Meseguer, Claude Marché & Xavier Urbain (2004): Proving termination of membership equational programs. In Nevin Heintze & Peter Sestoft, editors: Proceedings of the 2004 ACM SIGPLAN Workshop on Partial Evaluation and Semantics-based Program Manipulation, ACM, pp. 147–158, doi:10.1145/1014007.1014022.
- [5] Guillaume Feuillade & Thomas Genet (2003): *Reachability in Conditional Term Rewriting Systems*. Electronic Notes in Theoretical Computer Science 86(1), pp. 133–146, doi:10.1016/S1571-0661(04)80658-3.
- [6] Jürgen Giesl, Peter Schneider-Kamp & René Thiemann (2006): AProVE 1.2: Automatic Termination Proofs in the Dependency Pair Framework. In Ulrich Furbach & Natarajan Shankar, editors: Proceedings of the 3rd International Joint Conference on Automated Reasoning, Lecture Notes in Computer Science 4130, Springer, pp. 281–286, doi:10.1007/11814771_24.
- [7] Karl Gmeiner (2015): CoScart: Confluence Prover in Scala. In Ashish Tiwari & Takahito Aoto, editors: Proceedings of the 4th International Workshop on Confluence, p. 45.
- [8] Karl Gmeiner & Naoki Nishida (2014): Notes on Structure-Preserving Transformations of Conditional Term Rewrite Systems. In Manfred Schmidt-Schau
 ß, Masahiko Sakai, David Sabel & Yuki Chiba, editors: Proceedings of the first International Workshop on Rewriting Techniques for Program Transformations and Evaluation, OpenAccess Series in Informatics 40, Schloss Dagstuhl – Leibniz-Zentrum f
 ür Informatik, pp. 3–14, doi:10.4230/OASIcs.WPTE.2014.3.
- [9] Karl Gmeiner, Naoki Nishida & Bernhard Gramlich (2013): Proving Confluence of Conditional Term Rewriting Systems via Unravelings. In Nao Hirokawa & Vincent van Oostrom, editors: Proceedings of the 2nd International Workshop on Confluence, pp. 35–39.
- [10] Raúl Gutiérrez, Salvador Lucas & Patricio Reinoso (2016): A tool for the automatic generation of logical models of order-sorted first-order theories. In Alicia Villanueva, editor: Proceedings of the XVI Jornadas sobre Programación y Lenguages, pp. 215–230. Tool available at http://zenon.dsic.upv.es/ages/.
- [11] Manuel V. Hermenegildo & Francesca Rossi (1989): On the Correctness and Efficiency of Independent And-Parallelism in Logic Programs. In Ewing L. Lusk & Ross A. Overbeek, editors: Proceedings of the North American Conference on Logic Programming, MIT Press, pp. 369–389.
- [12] Jean-Marie Hullot (1980): Canonical Forms and Unification. In Wolfgang Bibel & Robert A. Kowalski, editors: Proceedings of the 5th Conference on Automated Deduction, Lecture Notes in Computer Science 87, Springer, pp. 318–334, doi:10.1007/3-540-10009-1_25.
- [13] Salvador Lucas (2017): A Semantic Approach to the Analysis of Rewriting-Based Systems. In Fabio Fioravanti & John P. Gallagher, editors: Pre-Proceedings of the 27th International Symposium on Logic-Based Program Synthesis and Transformation. CoRR, abs/1709.05095.
- [14] Salvador Lucas & Raúl Gutiérrez (2017): A Semantic Criterion for Proving Infeasibility in Conditional Rewriting. In Beniamino Accattoli & Bertram Felgenhauer, editors: Proceedings of the 6th International Workshop on Confluence, pp. 15–20.
- [15] Salvador Lucas, Claude Marché & José Meseguer (2005): Operational termination of conditional term rewriting systems. Information Processing Letters 95(4), pp. 446–453, doi:10.1016/j.ipl.2005.05.002.

- [16] Massimo Marchiori (1996): Unravelings and Ultra-properties. In Michael Hanus & Mario Rodríguez-Artalejo, editors: Proceedings of the 5th International Conference on Algebraic and Logic Programming, Lecture Notes in Computer Science 1139, Springer, pp. 107–121, doi:10.1007/3-540-61735-3_7.
- [17] Aart Middeldorp & Erik Hamoen (1994): Completeness Results for Basic Narrowing. Applicable Algebra in Engineering, Communication and Computing 5, pp. 213–253, doi:10.1007/BF01190830.
- [18] Naoki Nishida, Takayuki Kuroda, Makishi Yanagisawa & Karl Gmeiner (2015): *CO3: a COnverter for proving COnfluence of COnditional TRSs.* In Ashish Tiwari & Takahito Aoto, editors: *Proceedings of the 4th International Workshop on Confluence*, p. 42.
- [19] Naoki Nishida & Yuya Maeda (2018): Narrowing Trees for Syntactically Deterministic Conditional Term Rewriting Systems. In Hélène Kirchner, editor: Proceedings of the 3rd International Conference on Formal Structures for Computation and Deduction, Leibniz International Proceedings in Informatics 108, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, pp. 26:1–26:20, doi:10.4230/LIPIcs.FSCD.2018.26.
- [20] Naoki Nishida & Germán Vidal (2014): A Finite Representation of the Narrowing Space. In Gopal Gupta & Ricardo Peña, editors: Revised Selected Papers of the 23rd International Symposium on Logic-Based Program Synthesis and Transformation, Lecture Notes in Computer Science 8901, Springer, pp. 54–71, doi:10.1007/978-3-319-14125-1_4.
- [21] Naoki Nishida & Germán Vidal (2015): A framework for computing finite SLD trees. Journal of Logic and Algebraic Methods in Programming 84(2), pp. 197–217, doi:10.1016/j.jlamp.2014.11.006.
- [22] Enno Ohlebusch (2002): Advanced Topics in Term Rewriting. Springer, doi:10.1007/978-1-4757-3661-8.
- [23] Catuscia Palamidessi (1990): Algebraic Properties of Idempotent Substitutions. In Mike Paterson, editor: Proceedings of the 17th International Colloquium on Automata, Languages and Programming, Lecture Notes in Computer Science 443, Springer, pp. 386–399, doi:10.1007/BFb0032046.
- [24] Thomas Sternagel & Aart Middeldorp (2014): Conditional Confluence (System Description). In Gilles Dowek, editor: Proceedings of the Joint International Conference on Rewriting and Typed Lambda Calculi, Lecture Notes in Computer Science 8560, Springer, pp. 456–465, doi:10.1007/978-3-319-08918-8_31.
- [25] Taro Suzuki, Aart Middeldorp & Tetsuo Ida (1995): Level-Confluence of Conditional Rewrite Systems with Extra Variables in Right-Hand Sides. In Jieh Hsiang, editor: Proceedings of the 6th International Conference on Rewriting Techniques and Applications, Lecture Notes in Computer Science 914, Springer, pp. 179–193, doi:10.1007/3-540-59200-8_56.