Ground Canonical Rewrite Systems Revisited

Aart Middeldorp[1], Masahiko Sakai[2], and Sarah Winkler[3]

[1] University of Innsbruck, Innsbruck, Austria, aart.middeldorp@uibk.ac.at
[2] Nagoya University, Nagoya, Japan, sakai@i.nagoya-u.ac.jp
[3] Free University of Bozen-Bolzano, Bolzano, Italy, winkler@inf.unibz.it

Abstract

Systems of ground equations can always be transformed into equivalent canonical rewrite systems. Moreover, only finitely many distinct canonical rewrite systems exist for a given system of ground equations, a result proved by Snyder using congruence closure. Snyder also introduced a simple transformation to convert one canonical presentation into another one. In this paper we prove that this transformation is sound and complete, using standard rewrite techniques. We show that the transformation fails in the AC case.

1 Introduction

Congruence closure is an efficient technique to solve the word problem for systems of ground equations. Completion is a well-known technique for transforming systems of equations into equivalent canonical rewrite systems. It takes a reduction order as input and if it succeeds, the word problem is decidable by computing and comparing the unique normal forms of the two terms involved. For systems of ground equations, completion can be tamed such that it always terminates. Snyder [5] proved that the number of distinct canonical rewrite systems representing a given set of ground equations is at most 2^k where k is the number of equations. Furthermore, each of these canonical rewrite systems has the same number of rewrite rules. Finally, given one canonical rewrite system, all others can be obtained by a simple transformation.

This transformation is the topic of the paper. Using traditional rewrite techniques, which we briefly recall in Section 2, we prove in Section 3 that the transformation preserves canonicity and that it is complete in the sense that all canonical presentations of a given system of ground equations can be obtained from any of them by a finite number of transformation steps. In Section 4 we consider the extension of Snyder's transformation in the presence of associative and commutative operators. We conclude with some open questions.

2 Preliminaries

We assume familiarity with term rewriting but recall some important concepts and results in this preliminary section. An equational system (ES for short) is a set of equations between terms over a common signature. Throughout this paper we will consider finite ground TRSs. Every ground ES is also a TRS, and vice versa. A TRS is right-reduced if the right-hand sides of its rewrite rules are normal forms. It is left-reduced if every left-hand side of a rewrite rule is a normal form with respect to the other rules. A *reduced* TRS is both left-reduced and right-reduced. A TRS that is confluent, terminating and reduced is called *canonical*. Given a TRS $\mathcal{R}$, we denote the set of left-hand (right-hand) sides of its rules by $\mathsf{LHS}(\mathcal{R})$ ($\mathsf{RHS}(\mathcal{R})$). The set of its normal forms is denoted by $\mathsf{NF}(\mathcal{R})$.

Two TRSs $\mathcal{R}$ and $\mathcal{S}$ are *(conversion) equivalent* if $\leftrightarrow_{\mathcal{R}}^* = \leftrightarrow_{\mathcal{S}}^*$ and *normalization equivalent* if $\to_{\mathcal{R}}^! = \to_{\mathcal{S}}^!$. Here $t \to_{\mathcal{R}}^! u$ if both $t \to_{\mathcal{R}}^* u$ and $u \in \mathsf{NF}(\mathcal{R})$. A simple sufficient condition for normalization equivalence of canonical TRSs is stated in the following lemma, which is a special case of [1, Lemma 4.4(2)].

Lemma 1. *If* $\mathcal{R}$ *and* $\mathcal{S}$ *are canonical* TRSs *such that* $\mathsf{NF}(\mathcal{S}) \subseteq \mathsf{NF}(\mathcal{R})$ *and* $\to_{\mathcal{S}} \subseteq \leftrightarrow^*_{\mathcal{R}}$ *then* $\mathcal{R}$ *and* $\mathcal{S}$ *are normalization equivalent.* □

We conclude this preliminary section with two basic results that will be used in the sequel. We refer to [1] for modern (formalized) proofs of these results. The first one is a variation of a result by Métivier [3] and the second one is due to Snyder [5, Theorem 2.14].

Theorem 2.

1. *Normalization equivalent reduced* TRSs *are unique up to literal similarity.*
2. *Reduced ground* TRSs *are canonical.*

For ground TRSs, literal similarity amounts to equality, so normalization equivalent reduced ground TRSs are unique and equivalent canonical ground TRSs compatible with the same reduction order are identical.

3 Snyder's Transformation

Snyder [5, Theorem 3.18] showed how any set of ground equations $\mathcal{E}$ can be transformed into a special structure-sharing dag G from which an abstract relation T is extracted which represents a canonical ground TRS $\mathcal{R}$ equivalent to $\mathcal{E}$. The procedure runs in $O(n \log n)$ time, where n is the size of $\mathcal{E}$, though it may require more than $O(n \log n)$ time to actually obtain the TRS $\mathcal{R}$ from the dag. The method is complete in the sense that any canonical ground TRS $\mathcal{R}$ equivalent to $\mathcal{E}$ can be obtained in this way [5, Theorem 4.6]. This completeness result relies on the existence of a compatible well-founded order with the subterm property that is total on ground terms [5]. Furthermore, if $\mathcal{E}$ consists of k equations then there are at most 2^k equivalent canonical ground TRSs [5, Theorem 4.7]. This final result relies on the fact that any equivalent canonical TRS has at most k rewrite rules (which follows from ground completion, as remarked in [5, p. 424]).

Snyder mentions at the end of Section 4 that a single transformation is sufficient to generate all other equivalent canonical ground TRS from any given canonical TRS. This transformation is defined as follows:

$$\mathcal{R} \uplus \{\ell \to r\} \implies \mathcal{R}' \cup \{r \to \ell\} \tag{$\star$}$$

Here $\mathcal{R} \uplus \{\ell \to r\}$ is a canonical ground TRS such that r is not a proper subterm of ℓ, and $\mathcal{R}'$ is the ground TRS obtained from $\mathcal{R}$ by replacing every occurrence of r in the rewrite rules by ℓ. No proofs are provided, cf. [5, footnote 15].

Example 3. The collection of ground equations

$$\begin{array}{rcc} \mathsf{f}(\mathsf{a}) \approx \mathsf{g}(\mathsf{b},\mathsf{b}) & \mathsf{f}(\mathsf{f}(\mathsf{a})) \approx \mathsf{a} & \mathsf{f}(\mathsf{f}(\mathsf{f}(\mathsf{a}))) \approx \mathsf{a} \\ \mathsf{g}(\mathsf{b},\mathsf{h}(\mathsf{a})) \approx \mathsf{g}(\mathsf{b},\mathsf{b}) & \mathsf{h}(\mathsf{a}) \approx \mathsf{b} & \mathsf{i}(\mathsf{f}(\mathsf{a})) \approx \mathsf{c} \end{array}$$

admits six different canonical TRSs, which are connected by ($\star$) as depicted in Figure 1, where the labels of arrows indicate the rule that was flipped.

Below we present detailed proofs concerning the transformation ($\star$). We write $t\{r \mapsto \ell\}$ for the term obtained from t after replacing all subterms r by ℓ. The canonicity of $\mathcal{R}' \cup \{r \to \ell\}$ is relatively easy to prove.

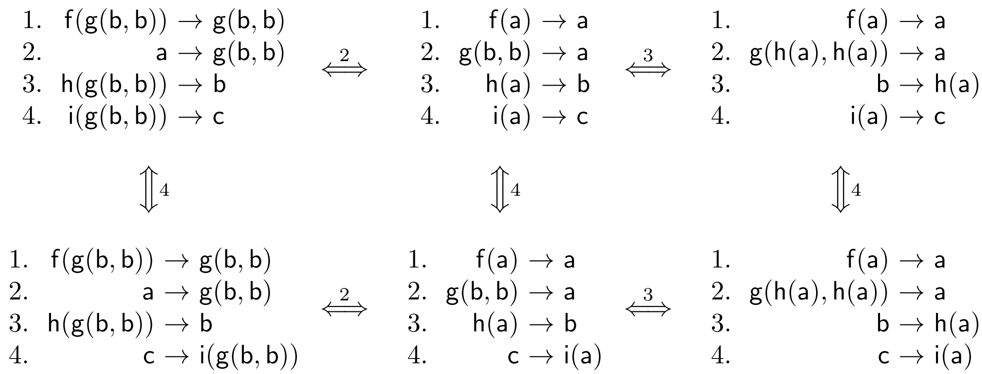


Figure 1: Six equivalent canononical TRSs.

Lemma 4. *The* TRS $\mathcal{R}' \cup \{r \to \ell\}$ *is canonical.*

Proof. We first show that $\mathcal{R}' \cup \{r \to \ell\}$ is right-reduced. Suppose to the contrary that a right-hand side t of a rule in $\mathcal{R}' \cup \{r \to \ell\}$ is reducible with a rule $u \to v \in \mathcal{R}' \cup \{r \to \ell\}$. Since we deal with ground TRSs, $t \trianglerighteq u$. Since right-hand sides of rules in $\mathcal{R}' \cup \{r \to \ell\}$ do not contain occurrences of r, $u \to v \in \mathcal{R}'$. Let $u' \to v'$ be the rule in $\mathcal{R}$ such that $u = u'\{r \mapsto \ell\}$ and $v = v'\{r \mapsto \ell\}$. Similarly, let t' be the right-hand side of a rule in $\mathcal{R}$ such that $t = t'\{r \mapsto \ell\}$. From $t \trianglerighteq u$ we infer $t' \trianglerighteq u'$, contradicting the fact that $\mathcal{R} \cup \{\ell \to r\}$ is right-reduced. Next we show that $\mathcal{R}' \cup \{r \to \ell\}$ is left-reduced. Let t be the left-hand side of a rewrite rule in $\mathcal{R}' \cup \{r \to \ell\}$. We distinguish two cases.

- Suppose $t = r \notin \mathsf{NF}(\mathcal{R}')$. So $r \trianglerighteq v$ for some left-hand side v of a rewrite rule in $\mathcal{R}'$. Since ℓ is not a proper subterm of r by assumption and $\ell = r$ is excluded by the right-reducedness of $\mathcal{R} \cup \{\ell \to r\}$, v contains no occurrences of ℓ. It follows that v is the left-hand side of a rewrite rule in $\mathcal{R}$, contradicting the right-reducedness of $\mathcal{R} \cup \{\ell \to r\}$.

- Suppose $t \to u \in \mathcal{R}'$ with $t \notin \mathsf{NF}((\mathcal{R}' \cup \{r \to \ell\}) \setminus \{t \to u\})$. Since t does not contain any occurrences of r, we have $t \notin \mathsf{NF}(\mathcal{R}' \setminus \{t \to u\})$. Let $v \to w$ be a rewrite rule in $\mathcal{R}' \setminus \{t \to u\}$ such that $t \trianglerighteq v$. Let t' and v' be the left-hand sides of rules in $\mathcal{R}$ such that $t = t'\{r \mapsto \ell\}$ and $v = v'\{r \mapsto \ell\}$. We have $t' \trianglerighteq v'$, contradicting left-reducedness of $\mathcal{R}$.

The proof is concluded by the canonicity of reduced ground TRSs, cf. Theorem 2 (2). □

We next show that every canonical presentation $\mathcal{S}$ of an ES $\mathcal{E}$ can be obtained from another canonical presentation $\mathcal{R}$ by a sequence of ($\star$) transformations.

Theorem 5. *Transformation ($\star$) is complete.*

Proof. Let $\mathcal{E}$ be an ES and $\mathcal{R}$ a canonical presentation of $\mathcal{E}$. For an arbitrary canonical representation $\mathcal{S}$ of $\mathcal{E}$, we prove by induction on $|\mathsf{RHS}(\mathcal{S}) \setminus \mathsf{NF}(\mathcal{R})|$ that $\mathcal{R}$ can be transformed into $\mathcal{S}$ by a sequence of ($\star$) transformations. First, note the following:

$$\mathsf{NF}(\mathcal{R}) \cap \mathsf{LHS}(\mathcal{S}) = \varnothing \implies \mathcal{R} = \mathcal{S} \qquad (\dagger)$$

This can be shown as follows. From the assumption we obtain $\mathsf{NF}(\mathcal{R}) \subseteq \mathsf{NF}(\mathcal{S})$ as $\mathsf{NF}(\mathcal{R})$ is closed under subterms. Conversion equivalence implies $\to_{\mathcal{R}} \subseteq \leftrightarrow^*_{\mathcal{S}}$ and $\mathcal{R}$ is terminating. Hence $\mathcal{R}$ and $\mathcal{S}$ are normalization equivalent by Lemma 1 and equal by Theorem 2 (1).

In the base case $\mathsf{RHS}(\mathcal{S}) \subseteq \mathsf{NF}(\mathcal{R})$. By (†), if $\mathcal{S} \neq \mathcal{R}$, there is some $u \to v \in \mathcal{S}$ such that $u \in \mathsf{NF}(\mathcal{R})$. However, $\mathsf{RHS}(\mathcal{S}) \subseteq \mathsf{NF}(\mathcal{R})$ implies $v \in \mathsf{NF}(\mathcal{R})$, which contradicts conversion equivalence as u and v are different convertible normal forms in $\mathcal{R}$.

In the induction step, we can assume by (†) that $\mathcal{S}$ contains a rule $u \to v$ such that $u \in \mathsf{NF}(\mathcal{R})$. Conversion equivalence and completeness of $\mathcal{R}$ imply $v \to^+_{\mathcal{R}} u$. In particular, v is not a proper subterm of u and hence we can apply transformation ($\star$) to $u \to v$ in $\mathcal{S}$. Let $\mathcal{S}'$ be the resulting TRS, which is canonical and contains $v \to u$. We compare $\mathsf{RHS}(\mathcal{S}) \setminus \mathsf{NF}(\mathcal{R})$ and $\mathsf{RHS}(\mathcal{S}') \setminus \mathsf{NF}(\mathcal{R})$. The former contains v as $u \to v \in \mathcal{S}$ and $v \notin \mathsf{NF}(\mathcal{R})$. This rule is replaced by $v \to u$ in $\mathcal{S}'$, which does not contribute to $\mathsf{RHS}(\mathcal{S}') \setminus \mathsf{NF}(\mathcal{R})$ as $u \in \mathsf{NF}(\mathcal{R})$. In addition, ($\star$) may replace right-hand sides $r[v] \in \mathsf{RHS}(\mathcal{S})$ by $r[u] \in \mathsf{RHS}(\mathcal{S}')$, but for all such terms $r[v]$, we have $r[v] \in \mathsf{RHS}(\mathcal{S}) \setminus \mathsf{NF}(\mathcal{R})$. Independent of whether or not $r[u] \in \mathsf{RHS}(\mathcal{S}') \setminus \mathsf{NF}(\mathcal{R})$ for some of the modified right-hand sides, we have $|\mathsf{RHS}(\mathcal{S}) \setminus \mathsf{NF}(\mathcal{R})| > |\mathsf{RHS}(\mathcal{S}') \setminus \mathsf{NF}(\mathcal{R})|$. By the induction hypothesis a sequence of ($\star$) transformations can turn $\mathcal{S}'$ into $\mathcal{R}$. □

4 Ground AC-Canonical Rewrite Systems

Marché [2, Theorem 3.1] proved that any AC canonical ground TRS for a finite set of ground equations with AC operators must be finite. The interesting proof relies on Dickson's Lemma. In the same paper, Marché presents a version of ground completion for the AC setting and a strategy that ensures termination [2, Theorem 4.3]. Unlike ground completion, in the AC setting critical pairs involving rules with the same AC symbol at the root of left-hand side need to be deduced. Unlike for AC completion, AC unification is not needed for computing these critical pairs. Ground AC completion relies on an AC simplification order which is AC total on ground terms. The existence of such an order was first shown by Narendran and Rusinowitch [4].

The next example shows that the transformation ($\star$) is unsuitable in an AC setting.

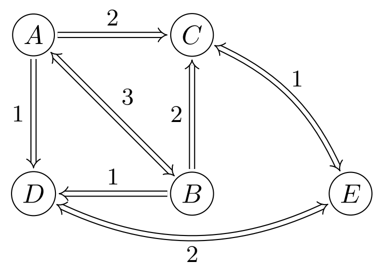
Example 6. The ground equations

$$\mathsf{f}(\mathsf{a},\mathsf{b}) \approx \mathsf{d} \qquad\qquad \mathsf{f}(\mathsf{b},\mathsf{c}) \approx \mathsf{e}$$

with AC symbol f admit five different AC canonical TRSs:

A	$\mathsf{f}(\mathsf{a},\mathsf{b}) \xrightarrow{1} \mathsf{d}$	$\mathsf{f}(\mathsf{b},\mathsf{c}) \xrightarrow{2} \mathsf{e}$	$\mathsf{f}(\mathsf{a},\mathsf{e}) \xrightarrow{3} \mathsf{f}(\mathsf{c},\mathsf{d})$
B	$\mathsf{f}(\mathsf{a},\mathsf{b}) \xrightarrow{1} \mathsf{d}$	$\mathsf{f}(\mathsf{b},\mathsf{c}) \xrightarrow{2} \mathsf{e}$	$\mathsf{f}(\mathsf{a},\mathsf{e}) \xleftarrow{3} \mathsf{f}(\mathsf{c},\mathsf{d})$
C	$\mathsf{f}(\mathsf{a},\mathsf{b}) \xrightarrow{1} \mathsf{d}$	$\mathsf{f}(\mathsf{b},\mathsf{c}) \xleftarrow{2} \mathsf{e}$	
D	$\mathsf{f}(\mathsf{a},\mathsf{b}) \xleftarrow{1} \mathsf{d}$	$\mathsf{f}(\mathsf{b},\mathsf{c}) \xrightarrow{2} \mathsf{e}$	
E	$\mathsf{f}(\mathsf{a},\mathsf{b}) \xleftarrow{1} \mathsf{d}$	$\mathsf{f}(\mathsf{b},\mathsf{c}) \xleftarrow{2} \mathsf{e}$	

If we apply ($\star$) to A by reversing rule 1 then rule 3 is first modified to $\mathsf{f}(\mathsf{a},\mathsf{e}) \approx \mathsf{f}(\mathsf{a},\mathsf{b},\mathsf{c})$ and subsequently deleted due to rule 2, resulting in D. (This cannot happen in the non-AC case.) Applying the AC version of ($\star$) systematically yields the following diagram:



Selecting rule 1 in D or rule 2 in C results in the TRS

$$F \qquad \mathsf{f}(\mathsf{a},\mathsf{b}) \xrightarrow{1} \mathsf{d} \qquad \mathsf{f}(\mathsf{b},\mathsf{c}) \xrightarrow{2} \mathsf{e}$$

which is not AC confluent. From F we obtain A and B by orienting the single *AC critical pair*.

The underlying problem is that reduced ground TRSs need not be AC confluent, necessitating the computation of AC critical pairs. In general, however, after applying ($\star$) and resolving AC critical pairs, new critical pairs may arise. Worse, reversing a rule might violate termination.

Example 7. The TRS $\mathcal{R}$ consisting of the two rules

$$\mathsf{f}(\mathsf{b},\mathsf{c}) \to \mathsf{f}(\mathsf{a},\mathsf{b}) \qquad \mathsf{f}(\mathsf{c},\mathsf{d}) \to \mathsf{f}(\mathsf{d},\mathsf{a})$$

with AC symbol f is AC canonical. Reversing the first rule results in the TRS $\mathcal{R}'$:

$$\mathsf{f}(\mathsf{a},\mathsf{b}) \to \mathsf{f}(\mathsf{b},\mathsf{c}) \qquad \mathsf{f}(\mathsf{c},\mathsf{d}) \to \mathsf{f}(\mathsf{d},\mathsf{a})$$

Transformation ($\star$) requires no further changes. However, AC termination is violated:

$$\mathsf{f}(\mathsf{a},\mathsf{b},\mathsf{d}) \to_{\mathcal{R}'/\mathsf{AC}} \mathsf{f}(\mathsf{b},\mathsf{c},\mathsf{d}) \to_{\mathcal{R}'/\mathsf{AC}} \mathsf{f}(\mathsf{a},\mathsf{b},\mathsf{d})$$

Despite the fact that the first AC completion procedures have been presented thirty years ago, many questions about even the simpler setting of ground AC completion are still open, for instance: Is the number of AC canonical presentations of an ES finite? If yes, how many such presentations are there? Can every equivalent AC canonical ground TRS be generated by modifying the AC termination order?

Acknowledgments. Part of this work was performed when the first author was employed at the Future Value Creation Research Center of Nagoya University, Japan.

References

[1] Nao Hirokawa, Aart Middeldorp, Christian Sternagel, and Sarah Winkler. Abstract completion, formalized. *Logical Methods in Computer Science*, 15(3):19:1–19:42, 2019.

[2] Claude Marché. On ground AC-completion. In *Proceedings of the 4th International Conference on Rewriting Techniques and Applications*, volume 488 of *Lecture Notes in Computer Science*, pages 411–422, 1991.

[3] Yves Métivier. About the rewriting systems produced by the Knuth–Bendix completion algorithm. *Information Processing Letters*, 16(1):31–34, 1983.

[4] Paliath Narendran and Michaël Rusinowitch. Any ground associative-commutative theory has a finite canonical system. In *Proceedings of the 4th International Conference on Rewriting Techniques and Applications*, volume 488 of *Lecture Notes in Computer Science*, pages 423–434, 1991.

[5] Wayne Snyder. A fast algorithm for generating reduced ground rewriting systems from a set of ground equations. *Journal of Symbolic Computation*, 15(4):415–450, 1993.