Normalization Equivalence of Rewrite Systems*

Nao Hirokawa[1], Aart Middeldorp[2] and Christian Sternagel[2]

[1] JAIST, Japan
hirokawa@jaist.ac.jp
[2] University of Innsbruck, Austria
{aart.middeldorp|christian.sternagel}@uibk.ac.at

Abstract

Métivier (1983) proved that every confluent and terminating rewrite system can be transformed into an equivalent canonical rewrite system. He also proved that equivalent canonical rewrite systems which are compatible with the same reduction order are unique up to variable renaming. In this note we present simple and formalized proofs of these results. The latter result is generalized to the uniqueness of *normalization equivalent* reduced rewrite systems.

1 Introduction

Consider the TRS $\mathcal{R}$ of combinatory logic with equality test, studied by Klop [3]:

$$\mathsf{S}xyz \to xz(yz) \qquad \mathsf{K}xy \to x \qquad \mathsf{I}x \to x \qquad \mathsf{D}xx \to \mathsf{E}$$

The TRS $\mathcal{R}$ is reduced, but neither terminating nor confluent. One might ask: *is there another reduced TRS $\mathcal{S}$ that computes the same normal forms for every starting term?* We refer to this property as *normalization equivalence* of two TRSs. According to the main result of this note, it turns out that $\mathcal{R}$ is unique up to variable renaming.

In the next section normalization equivalence is studied in an abstract setting. The concrete results on term rewrite systems are presented in Section 3. Throughout this note, we assume familiarity with basic notions and terminology of term rewriting.

All the proofs that are presented in the following have been formalized as part of IsaFoR[1] (see theory `Normalization_Equivalence`).

2 Abstract Normalization Equivalence

First, we introduce the two notions of equivalence that will be studied in this note.

Definition 2.1. Two ARSs $\mathcal{A}$ and $\mathcal{B}$ are *(conversion) equivalent* if $\leftrightarrow^*_{\mathcal{A}} = \leftrightarrow^*_{\mathcal{B}}$. If $\to^!_{\mathcal{A}} = \to^!_{\mathcal{B}}$ we say that $\mathcal{A}$ and $\mathcal{B}$ are *normalization equivalent*.

The following example shows that the two equivalence notions defined above are different.

Example 2.2. The ARSs

$$\mathcal{A}_1: \quad \mathsf{a} \longrightarrow \mathsf{b} \qquad\qquad \mathcal{B}_1: \quad \mathsf{a} \longleftarrow \mathsf{b}$$

*Supported by JSPS KAKENHI Grant Number 25730004 and the Austrian Science Fund (FWF) projects I963 and J3202.

[1] `http://cl-informatik.uibk.ac.at/software/ceta/`

are conversion equivalent but not normalization equivalent. The ARSs

are normalization equivalent but not conversion equivalent.

The easy proof (by induction on the length of conversions) of the following result is omitted.

Lemma 2.3. *Normalization equivalent terminating ARSs are equivalent.* □

Note that the termination assumption can be weakened to weak normalization. However, the present version suffices to prove the following lemma that we employ in our proof of Métivier's transformation result (Theorem 3.7).

Lemma 2.4. *Let $\mathcal{A}$ and $\mathcal{B}$ be ARSs such that $\to_{\mathcal{B}} \subseteq \to^{+}_{\mathcal{A}}$ and $\mathsf{NF}(\mathcal{B}) \subseteq \mathsf{NF}(\mathcal{A})$. If $\mathcal{A}$ is complete then $\mathcal{B}$ is complete and normalization equivalent to $\mathcal{A}$.*

Proof. From the inclusion $\to_{\mathcal{B}} \subseteq \to^{+}_{\mathcal{A}}$ we infer that $\mathcal{B}$ is terminating. Moreover, $\to^{*}_{\mathcal{B}} \subseteq \to^{*}_{\mathcal{A}}$ and, since $\mathsf{NF}(\mathcal{B}) \subseteq \mathsf{NF}(\mathcal{A})$, also $\to^{!}_{\mathcal{B}} \subseteq \to^{!}_{\mathcal{A}}$. For the reverse inclusion we reason as follows. Let $a \to^{!}_{\mathcal{A}} b$. Because $\mathcal{B}$ is terminating, $a \to^{!}_{\mathcal{B}} c$ for some $c \in \mathsf{NF}(\mathcal{B})$. So $a \to^{!}_{\mathcal{A}} c$ and thus $b = c$ from the confluence of $\mathcal{A}$. It follows that $\mathcal{A}$ and $\mathcal{B}$ are normalization equivalent. It remains to show that $\mathcal{B}$ is locally confluent. This follows from the sequence of inclusions

$$_{\mathcal{B}}{\leftarrow} \cdot \to_{\mathcal{B}} \;\subseteq\; {}^{+}_{\mathcal{A}}{\leftarrow} \cdot \to^{+}_{\mathcal{A}} \;\subseteq\; \to^{*}_{\mathcal{A}} \cdot {}^{*}_{\mathcal{A}}{\leftarrow} \;\subseteq\; \to^{!}_{\mathcal{A}} \cdot {}^{!}_{\mathcal{A}}{\leftarrow} \;\subseteq\; \to^{!}_{\mathcal{B}} \cdot {}^{!}_{\mathcal{B}}{\leftarrow}$$

where we use the inclusion $\to_{\mathcal{B}} \subseteq \to^{+}_{\mathcal{A}}$, the confluence of $\mathcal{A}$, the termination of $\mathcal{A}$, and the normalization equivalence of $\mathcal{A}$ and $\mathcal{B}$. □

In the above lemma, completeness can be weakened to semi-completeness (i.e., the combination of confluence and weak normalization), which is not true for Theorem 3.7 as shown by Gramlich [1]. Again, the present version suffices for our purposes.

3 Normalization Equivalence

In this section we study normalization equivalence for TRSs.

Definition 3.1. A *variable substitution* is a substitution from $\mathcal{V}$ to $\mathcal{V}$. A *renaming* is a bijective variable substitution. A term s is a *variant* of a term t if $s = t\sigma$ for some renaming σ. If $\ell \to r$ is a rewrite rule and σ is a renaming then the rewrite rule $\ell\sigma \to r\sigma$ is a variant of $\ell \to r$. A TRS is said to be *variant-free* if it does not contain rewrite rules that are variants of each other.

TRSs are usually assumed to be variant-free. We make the same assumption, but see Example 3.6 below.

Given terms s and t, we write $s \doteq t$ if $s\sigma = t$ and $s = t\tau$ for some substitutions σ and τ. The following result is folklore; the proof has recently been formalized [2].

Lemma 3.2. *Two terms s and t are variants if and only if $s \doteq t$.* □

Definition 3.3. Two TRSs $\mathcal{R}_1$ and $\mathcal{R}_2$ over the same signature $\mathcal{F}$ are called *literally similar*, denoted by $\mathcal{R}_1 \doteq \mathcal{R}_2$, if every rewrite rule in $\mathcal{R}_1$ has a variant in $\mathcal{R}_2$ and vice-versa.

Definition 3.4. A TRS $\mathcal{R}$ is *left-reduced* if $\ell \in \mathsf{NF}(\mathcal{R} \setminus \{\ell \to r\})$ for every rewrite rule $\ell \to r$ in $\mathcal{R}$. We say that $\mathcal{R}$ is *right-reduced* if $r \in \mathsf{NF}(\mathcal{R})$ for every rewrite rule $\ell \to r$ in $\mathcal{R}$. A *reduced* TRS is left- and right-reduced. A reduced complete TRS is called *canonical*.

Theorem 3.7 below states that we can always eliminate redundancy in a complete TRS. This is achieved by the two-stage transformation defined below.

Definition 3.5. Given a complete TRS $\mathcal{R}$, the TRSs $\dot{\mathcal{R}}$ and $\ddot{\mathcal{R}}$ are defined as follows:

$$\dot{\mathcal{R}} = \{\ell \to r{\downarrow_{\mathcal{R}}} \mid \ell \to r \in \mathcal{R}\}$$
$$\ddot{\mathcal{R}} = \{\ell \to r \in \dot{\mathcal{R}} \mid \ell \in \mathsf{NF}(\dot{\mathcal{R}} \setminus \{\ell \to r\})\}$$

The TRS $\dot{\mathcal{R}}$ is obtained from $\mathcal{R}$ by normalizing the right-hand sides. To obtain $\ddot{\mathcal{R}}$ we remove the rules of $\dot{\mathcal{R}}$ whose left-hand sides are reducible with another rule of $\dot{\mathcal{R}}$.

Example 3.6. Consider the TRS $\mathcal{R}_1$ consisting of the four rules

$$\mathsf{f}(x) \to \mathsf{a} \qquad \mathsf{f}(y) \to \mathsf{b} \qquad \mathsf{a} \to \mathsf{c} \qquad \mathsf{b} \to \mathsf{c}$$

Then the first transformation yields $\dot{\mathcal{R}}_1$

$$\mathsf{f}(x) \to \mathsf{c} \qquad \mathsf{f}(y) \to \mathsf{c} \qquad \mathsf{a} \to \mathsf{c} \qquad \mathsf{b} \to \mathsf{c}$$

and the second one $\ddot{\mathcal{R}}_1$

$$\mathsf{a} \to \mathsf{c} \qquad \mathsf{b} \to \mathsf{c}$$

Note that $\ddot{\mathcal{R}}_1$ is *not* equivalent to $\mathcal{R}_1$. This is caused by the fact that the result of the first transformation is no longer variant-free.

The proof of the following theorem depends on the implicit assumption that TRSs are always variant-free. However, even for variant-free $\mathcal{R}$, $\dot{\mathcal{R}}$ does not necessarily have this property (as shown by Example 3.6 above). Thus, in our formalization, we explicitly remove variants of rules as part of the $\dot{\mathcal{R}}$ transformation.

Theorem 3.7. *If $\mathcal{R}$ is a complete TRS then $\ddot{\mathcal{R}}$ is a normalization and conversion equivalent canonical TRS.*

The proof by Métivier [4, Theorem 7] is hard to reconstruct. The proof in [5, Exercise 7.4.7] involves 13 steps with lots of redundancy. The proof below uses induction on the well-founded encompassment order $\rhd$ and has been formalized. Since subsumption as well as encompassment have not been part of IsaFoR before, we had to amend this situation. See theory `Encompassment` for details.

Proof. Let $\mathcal{R}$ be a complete TRS. The inclusions $\ddot{\mathcal{R}} \subseteq \dot{\mathcal{R}} \subseteq \to_{\mathcal{R}}^{+}$ are obvious from the definitions. Since $\mathcal{R}$ and $\dot{\mathcal{R}}$ have the same left-hand sides, their normal forms coincide. We show that $\mathsf{NF}(\ddot{\mathcal{R}}) \subseteq \mathsf{NF}(\dot{\mathcal{R}})$. To this end we show that $\ell \notin \mathsf{NF}(\ddot{\mathcal{R}})$ whenever $\ell \to r \in \dot{\mathcal{R}}$ by induction on ℓ with respect to the well-founded order $\rhd$. If $\ell \to r \in \ddot{\mathcal{R}}$ then $\ell \notin \mathsf{NF}(\ddot{\mathcal{R}})$ trivially holds. So suppose $\ell \to r \notin \ddot{\mathcal{R}}$. By definition of $\ddot{\mathcal{R}}$, $\ell \notin \mathsf{NF}(\dot{\mathcal{R}} \setminus \{\ell \to r\})$. So there exists a rewrite rule $\ell' \to r' \in \dot{\mathcal{R}}$ different from $\ell \to r$ such that $\ell \trianglerighteq \ell'$. We distinguish two cases.

- If $\ell \rhd \ell'$ then we obtain $\ell' \notin \mathsf{NF}(\ddot{\mathcal{R}})$ from the induction hypothesis and hence $\ell \notin \mathsf{NF}(\ddot{\mathcal{R}})$ as desired.

- If $\ell \doteq \ell'$ then by Lemma 3.2 there exists a renaming σ such that $\ell = \ell'\sigma$. Since $\dot{\mathcal{R}}$ is right-reduced by construction, r and r' are normal forms of $\dot{\mathcal{R}}$. The same holds for $r'\sigma$ because normal forms are closed under renaming. We have $r \;{}_{\dot{\mathcal{R}}}\!\leftarrow \ell = \ell'\sigma \rightarrow_{\dot{\mathcal{R}}} r'\sigma$. Since $\dot{\mathcal{R}}$ is confluent as a consequence of Lemma 2.4, $r = r'\sigma$. Hence $\ell' \rightarrow r'$ is a variant of $\ell \rightarrow r$, contradicting the assumption that TRSs are variant-free.

From Lemma 2.4 we infer that the TRSs $\dot{\mathcal{R}}$ and $\ddot{\mathcal{R}}$ are complete and normalization equivalent to $\mathcal{R}$. The TRS $\ddot{\mathcal{R}}$ is right-reduced because $\ddot{\mathcal{R}} \subseteq \dot{\mathcal{R}}$ and $\dot{\mathcal{R}}$ is right-reduced. From $\mathsf{NF}(\ddot{\mathcal{R}}) = \mathsf{NF}(\dot{\mathcal{R}})$ we easily infer that $\ddot{\mathcal{R}}$ is left-reduced. It follows that $\ddot{\mathcal{R}}$ is canonical. It remains to show that $\ddot{\mathcal{R}}$ is not only normalization equivalent but also (conversion) equivalent to $\mathcal{R}$. This is an immediate consequence of Lemma 2.3. □

For our next result we need the following technical lemma.

Lemma 3.8. *Let $\mathcal{R}$ be a right-reduced TRS and let s be a reducible term which is minimal with respect to $\rhd$. If $s \rightarrow^+_{\mathcal{R}} t$ then $s \rightarrow t$ is a variant of a rule in $\mathcal{R}$*

Proof. Let $\ell \rightarrow r$ be the rewrite rule that is used in the first step from s to t. So $s \unrhd \ell$. By assumption, $s \rhd \ell$ does not hold and thus $s \doteq \ell$. According to Lemma 3.2 there exists a renaming σ such that $s = \ell\sigma$. We have $s \rightarrow_{\mathcal{R}} r\sigma \rightarrow^*_{\mathcal{R}} t$. Because $\mathcal{R}$ is right-reduced, $r \in \mathsf{NF}(\mathcal{R})$. Since normal forms are closed under renaming, also $r\sigma \in \mathsf{NF}(\mathcal{R})$ and thus $r\sigma = t$. It follows that $s \rightarrow t$ is a variant of $\ell \rightarrow r$. □

In our formalization, the above proof is the first spot where we actually need that $\mathcal{R}$ satisfies the variable condition (more precisely, right-hand sides of rules do not introduce fresh variables).

The next result is the main result of this note.

Theorem 3.9. *Normalization equivalent reduced TRSs are unique up to literal similarity.*

Proof. Let $\mathcal{R}$ and $\mathcal{S}$ be normalization equivalent reduced TRSs. Suppose $\ell \rightarrow r \in \mathcal{R}$. Because $\mathcal{R}$ is right-reduced, $r \in \mathsf{NF}(\mathcal{R})$ and thus $\ell \neq r$. Hence $\ell \rightarrow^+_{\mathcal{S}} r$ by normalization equivalence. Because $\mathcal{R}$ is left-reduced, ℓ is a minimal (with respect to $\rhd$) $\mathcal{R}$-reducible term. Another application of normalization equivalence yields that ℓ is minimal $\mathcal{S}$-reducible. Hence $\ell \rightarrow r$ is a variant of a rule in $\mathcal{S}$ by Lemma 3.8. □

We show that the corresponding result of Métivier [4, Theorem 8] is an easy consequence of Theorem 3.9. Here a TRS $\mathcal{R}$ is said to be compatible with a reduction order $>$ if $\ell > r$ for every rewrite rule $\ell \rightarrow r$ of $\mathcal{R}$.

Theorem 3.10. *Let $\mathcal{R}$ and $\mathcal{S}$ be equivalent canonical TRSs. If $\mathcal{R}$ and $\mathcal{S}$ are compatible with the same reduction order then $\mathcal{R} \doteq \mathcal{S}$.*

Proof. Suppose $\mathcal{R}$ and $\mathcal{S}$ are compatible with the reduction order $>$. We show that $\rightarrow^!_{\mathcal{R}} \subseteq \rightarrow^!_{\mathcal{S}}$. Let $s \rightarrow^!_{\mathcal{R}} t$. We show that $t \in \mathsf{NF}(\mathcal{S})$. Let u be the unique $\mathcal{S}$-normal form of t. We have $t \rightarrow^!_{\mathcal{S}} u$ and thus $t \leftrightarrow^*_{\mathcal{R}} u$ because $\mathcal{R}$ and $\mathcal{S}$ are equivalent. Since $t \in \mathsf{NF}(\mathcal{R})$, we have $u \rightarrow^!_{\mathcal{R}} t$. If $t \neq u$ then both $t > u$ (as $t \rightarrow^!_{\mathcal{S}} u$) and $u > t$ (as $u \rightarrow^!_{\mathcal{R}} t$), which is impossible. Hence $t = u$ and thus $t \in \mathsf{NF}(\mathcal{S})$. Together with $s \leftrightarrow^*_{\mathcal{S}} t$, which follows from the equivalence of $\mathcal{R}$ and $\mathcal{S}$, we conclude that $s \rightarrow^!_{\mathcal{S}} t$. We obtain $\rightarrow^!_{\mathcal{S}} \subseteq \rightarrow^!_{\mathcal{R}}$ by symmetry. Hence $\mathcal{R}$ and $\mathcal{S}$ are normalization equivalent and the result follows from Theorem 3.9. □

References

[1] B. Gramlich. On interreduction of semi-complete term rewriting systems. *Theoretical Computer Science*, 258(1-2):435–451, 2001. doi:10.1016/S0304-3975(00)00030-X.

[2] N. Hirokawa, A. Middeldorp, and C. Sternagel. A new and formalized proof of abstract completion. In *Proc. 5th International Conference on Interactive Theorem Proving*, volume 8558 of *Lecture Notes in Computer Science*, 2014. To appear.

[3] J.W. Klop. *Combinatory Reduction Systems.* PhD thesis, Utrecht University, 1980.

[4] Y. Métivier. About the rewriting systems produced by the Knuth-Bendix completion algorithm. *Information Processing Letters*, 16(1):31–34, 1983. doi:10.1016/0020-0190(83)90009-1.

[5] Terese. *Term Rewriting Systems*, volume 55 of *Cambridge Tracts in Theoretical Computer Science.* Cambridge University Press, 2003.