Available online at www.sciencedirect.com

SCIENCE DIRECT®

Information Processing Letters 91 (2004) 141–146

Information Processing Letters

www.elsevier.com/locate/ipl

On confluence and residuals in Cauchy convergent transfinite rewriting

Jakob Grue Simonsen

Department of Computer Science, University of Copenhagen (DIKU), Universitetsparken 1, 2100 Copenhagen Ø, Denmark

Received 5 May 2003; received in revised form 7 January 2004

Available online 10 May 2004

Communicated by A. Tarlecki

Abstract

We show that non-collapsing orthogonal term rewriting systems do not have the transfinite Church–Rosser property in the setting of Cauchy convergence. In addition, we show that for (a transfinite version of) the Parallel Moves Lemma to hold, any definition of residual for Cauchy convergent rewriting must *either* part with a number of fundamental properties enjoyed by rewriting systems in the finitary and strongly convergent settings, *or* fail to hold for very simple rewriting systems.

Keywords: Transfinite term rewriting; Infinitary rewriting; Cauchy convergence; Church–Rosser property; Programming calculi

Is there no limit?
— Job 16:3

No, no!
No, no, no, no!
No, no, no, no!
No, no, there's no limit!
— 2 Unlimited: "No Limit" (1993)

1. Introduction

Term rewriting [2,7,1,9] is a popular class of models well-suited for analysis of programming languages that has been extended to encompass infinite terms and infinite rewrite sequences in order to reason about, e.g., lazy data structures and non-terminating computations. In this setting of *transfinite rewriting*, many of the standard techniques of the field, such as orthogonality and residual theory, have been redeveloped in landmark papers [3,5]; see [4] for an updated introduction.

In the subfield of *strongly convergent* transfinite rewriting, confluence properties for orthogonal systems have been the object of extensive research [5,6,4], since the more general setting of *Cauchy convergent rewriting* did not readily yield useful results [3]. In [5] it was conjectured that the basic confluence results from strongly convergent rewriting could be lifted to Cauchy convergent rewriting.

In the present paper we note two negative results:

- First, we provide a counterexample to the aforementioned conjecture, demonstrating that the special case of the Strip Lemma fails, even for

E-mail address: simonsen@diku.dk (J.G. Simonsen).

0020-0190/$ – see front matter
doi:10.1016/j.ipl.2004.03.018

non-collapsing systems. The failure of the Strip Lemma entails in turn that it is impossible to have a workable "Transfinite Parallel Moves Lemma" for *all* orthogonal systems.
- Second, we exhibit examples showing that developing a theory of *residuals*—the traditional means of proving orthogonal systems confluent—faces further obstacles than those already identified in previous work on transfinite rewriting. In particular, we demonstrate that any definition of residual allowing for a Transfinite Parallel Moves Lemma to hold for a set, A, of transfinite rewriting systems, will *either* have to part with a number of the properties enjoyed by residual relations in finitary and strongly convergent rewriting, *or A* will fail to contain certain very simple orthogonal systems.

Our results provide negative answers to Problems 18 and 47 on the List of Open Problems in Rewriting (http://www.lsv.ens-cachan.fr/rtaloop/). In both cases, however, our results only amount to counterexamples to the *general* problem, and thus raise the question of whether desirable results may hold for suitably restricted systems; Section 4 shows that this may be an arduous task.

2. Preliminaries

We assume the reader to be familiar with the general theory of (first-order) term rewriting; ample introductions are [2,7,1] and Chapter 2 of the recent [9]. In addition, we assume at least passing familiarity with the basics of ordinal theory, corresponding to the level of [8]; the Greek letter ω denotes the least infinite ordinal. The definitions of transfinite rewrite systems of this paper are those of [4], with only minor differences in typography and nomenclature.

Positions in (finite) terms are elements of $\{1, 2, \ldots\}^*$ defined in the usual way. The subterm of term s at position p is denoted $s|_p$. If $\mathbf{f}$ is a unary function symbol and $k \in \omega$, we denote by $\mathbf{f}^k(s)$ k successive applications of $\mathbf{f}$ to the term s; we extend the notation to include $\mathbf{f}^\omega$ with the obvious meaning. Finally, we presuppose a denumerably infinite set of variables, χ, and a "Hilbert Hotel" method of renaming variables if the need for fresh variables arises.

Definition 1. Let $\mathrm{Ter}(\Sigma)$ be the set of (finite) terms over the (not necessarily finite) signature Σ (with variable set χ). Define the metric $d : \mathrm{Ter}(\Sigma) \times \mathrm{Ter}(\Sigma) \to [0; 1]$ by $d(t, t') \triangleq 0$ if $t = t'$, and $d(t, t') = 2^{-k}$ otherwise, where k is the length of the shortest position at which t and t' differ. The metric completion of the metric space $(\mathrm{Ter}(\Sigma), d)$, denoted $\mathrm{Ter}^\infty(\Sigma)$, is called the *set of finite and infinite terms over* Σ.

Definition 2. An *infinitary rewrite rule* is a pair $\mathbf{l} \to \mathbf{r}$ where $\mathbf{l} \in \mathrm{Ter}(\Sigma)$ and $\mathbf{r} \in \mathrm{Ter}^\infty(\Sigma)$ such that $\mathbf{l}$ is not a variable and every variable in $\mathbf{r}$ occurs in $\mathbf{l}$. An *infinitary term rewriting system*, denoted iTRS, is a pair $\mathcal{R} \triangleq (\Sigma, R)$, consisting of an signature Σ and a set of infinitary rewrite rules R.

The usual definitions of redex, rewrite step, left- and right-linearity, orthogonality, duplication, developments, etc., carry over in the immediate fashion from the finitary setting. Whenever a redex is contracted, we usually underline its head-symbol. If particular attention is to be paid to a redex (that is not necessarily contracted), we usually overline it.

Definition 3. Let a be an ordinal. A *derivation of length* α is a net of rewrite steps $(s_\beta \to s_{\beta+1})_{\beta<\alpha}$. In the step $s_\beta \to s_{\beta+1}$, assume that the redex contracted is at position u_β of s_β; the *depth*, d_β, of the redex is the length of u_β.

The derivation is called *Cauchy convergent* (or *weakly convergent*) if, for every limit ordinal $\lambda \leqslant \alpha$, the distance $d(s_\beta, s_\lambda)$ tends to 0 as β approaches λ from below. It is called *strongly convergent* if it is Cauchy convergent and, in addition, d_β tends to infinity as β approaches λ from below.

If $(s_\beta \to s_{\beta+1})_{\beta<\alpha}$ is convergent with limit t and $s_0 = s$, we write $s \to^\alpha t$, and say that t is a *derivative* of s. When the length of the derivation is bounded above by γ, we shall occasionally write $s \to^{\leqslant\gamma} t$. When the length of the derivation is unimportant, we write $s \twoheadrightarrow t$.

Strong convergence intuitively means that as a derivation approaches limit ordinal length, redexes are contracted "deeper and deeper". This seemingly innocuous feature is the crux of the proofs of virtually all interesting results pertaining to (orthogonal) iTRS;

one of the primary consequences of requiring strong convergence is that it is sufficient to consider derivations of length $\leqslant \omega$:

Lemma 4 (Compression Lemma). *Let $s \twoheadrightarrow t$ be a strongly convergent derivation in a left-linear iTRS. Then, there exists a strongly convergent derivation $s \rightarrow^{\leqslant \omega} t$.*

Definition 5. An iTRS is called *confluent* (or *transfinitely Church–Rosser*) if, for all $s, t, t' \in \mathrm{Ter}^{\infty}(\Sigma)$ such that $s \twoheadrightarrow t$ and $s \twoheadrightarrow t'$, there is an s' with $t \twoheadrightarrow s'$ and $t' \twoheadrightarrow s'$. If all of the involved derivations are *strongly* convergent, we say that the iTRS is *strongly confluent*.

Definition 6. A rule is called *collapsing* if its right-hand side is a variable. A redex is collapsing if the rule of the redex is collapsing. A term, s, is *hypercollapsing* if all derivatives of s have a derivative with a collapsing redex at the root.

An orthogonal iTRS is said to be *almost-non-collapsing* if it contains at most one collapsing rule $\mathbf{l} \to x$ and the only variable in $\mathbf{l}$ is x.

If a system is *not* almost-non-collapsing, it is easy to give counterexamples to confluence in both the strongly and Cauchy convergent settings [3,5]. The following positive results are due to Kennaway et al. [5] (see also [4]):

Theorem 7. *An orthogonal iTRS is strongly confluent iff it is almost non-collapsing.*

Theorem 8. *An orthogonal iTRS is strongly confluent modulo identification of hypercollapsing subterms.*

As a final preparation, we give the following auxiliary definition:

Definition 9. Let P be a predicate on $\mathrm{Ter}^{\infty}(\Sigma)$. A derivation $s \rightarrow^{\alpha} t$ is said to have property P α-frequently if, for each ordinal β with $\beta < \alpha$, there exists a γ with $\beta < \gamma < \alpha$ such that s_{γ} has property P.

3. Failure of confluence in Cauchy convergent rewriting

Consider the non-collapsing orthogonal iTRS R consisting of the following infinite set of rules:

$$\mathbf{a} \to \mathbf{b}$$
$$\mathbf{f}\big(\mathbf{g}^{k}(\mathbf{c}), x, y\big) \to \mathbf{f}\big(\mathbf{g}^{k+1}(\mathbf{c}), y, y\big) \quad \text{for } k \in \omega \text{ even,}$$
$$\mathbf{f}\big(\mathbf{g}^{k}(\mathbf{c}), x, y\big) \to \mathbf{f}\big(\mathbf{g}^{k+1}(\mathbf{c}), \mathbf{a}, y\big) \quad \text{for } k \in \omega \text{ odd.}$$

Observe that R is orthogonal—overlaps between rules of the form $\mathbf{f}(\cdot,\cdot,\cdot)$ are prevented by the different number of $\mathbf{g}$s in each rule. Ponder the derivation S of length ω:

$$\underline{\mathbf{f}}(\mathbf{c}, \mathbf{a}, \mathbf{a}) \to \underline{\mathbf{f}}\big(\mathbf{g}(\mathbf{c}), \mathbf{a}, \mathbf{a}\big) \to \underline{\mathbf{f}}\big(\mathbf{g}(\mathbf{g}(\mathbf{c})), \mathbf{a}, \mathbf{a}\big) \to \cdots$$

which is clearly Cauchy convergent with limit $\mathbf{f}(\mathbf{g}^{\omega}, \mathbf{a}, \mathbf{a})$, and note that contracting the overlined redex in $\mathbf{f}(\mathbf{c}, \mathbf{a}, \bar{\mathbf{a}})$ yields the term $\mathbf{f}(\mathbf{c}, \mathbf{a}, \mathbf{b})$.

Proposition 10. *$\mathbf{f}(\mathbf{g}^{\omega}, \mathbf{a}, \mathbf{a})$ and $\mathbf{f}(\mathbf{c}, \mathbf{a}, \mathbf{b})$ have no common derivative, i.e., R is not confluent.*

Proof. Observe that any derivative of $\mathbf{f}(\mathbf{g}^{\omega}, \mathbf{a}, \mathbf{a})$ must be on the form $\mathbf{f}(\mathbf{g}^{\omega}, \cdot, \cdot)$. We show that there is no convergent derivation from $\mathbf{f}(\mathbf{c}, \mathbf{a}, \mathbf{b})$ to a term on this form. Assume that β is the least ordinal such that a convergent derivation $S : \mathbf{f}(\mathbf{c}, \mathbf{a}, \mathbf{b}) \rightarrow^{\beta} w$ exists where w is of the form $\mathbf{f}(\mathbf{g}^{\omega}, \cdot, \cdot)$. Since β is the least such ordinal, there will β-frequently be a rewrite step at the root to change the $\mathbf{c}$ into $\mathbf{g}^{\omega}$. If $\mathbf{f}(\mathbf{c}, \mathbf{a}, \mathbf{b}) \rightarrow^{\alpha} t \xrightarrow{u} t'$ where the redex u is contracted at the root of t, then t is of the form $\mathbf{f}(\mathbf{g}^{k}(\mathbf{c}), \cdot, \cdot)$ and t' of the form $\mathbf{f}(\mathbf{g}^{k+1}(\mathbf{c}), \cdot, \cdot)$ for some $k \in \omega$. The steps performed at the root must thus be alternating applications of rules from

$$\big\{\mathbf{f}\big(\mathbf{g}^{k}(\mathbf{c}), x, y\big) \to \mathbf{f}\big(\mathbf{g}^{k+1}(\mathbf{c}), y, y\big) \colon k \in \omega \text{ even}\big\}$$

and

$$\big\{\mathbf{f}\big(\mathbf{g}^{k}(\mathbf{c}), x, y\big) \to \mathbf{f}\big(\mathbf{g}^{k+1}(\mathbf{c}), \mathbf{a}, y\big) \colon k \in \omega \text{ odd}\big\}.$$

Hence, in $S : \mathbf{f}(\mathbf{c}, \mathbf{a}, \mathbf{b}) \rightarrow^{\beta} w$, there will β-frequently be terms from $\{s\colon s|_2 = \mathbf{a}\}$. Inspection of the rules reveals that if $\mathbf{f}(\mathbf{c}, \mathbf{a}, \mathbf{b}) \rightarrow^{\beta} t$, then t must be of the form $\mathbf{f}(\cdot, \cdot, \mathbf{b})$; this entails that S will also β-frequently contain terms from both $\{s\colon s|_2 = \mathbf{b}\}$ and $\{s\colon s|_3 = \mathbf{b}\}$. Thus, there will β-frequently be an a and β-frequently be a $\mathbf{b}$ at position 2 of the terms of S, whence

$\mathbf{f}(\mathbf{c}, \mathbf{a}, \mathbf{b}) \to^{\beta} w$ cannot be convergent, contradicting the assumptions. □

Thus, even the special case of the Strip Lemma fails to hold for non-collapsing iTRS in the presence of Cauchy convergence. There are several quirks to the iTRS R: it is not right-linear, the right-hand sides of rules are not normal forms, and there is no bound on the depth of left-hand sides of rules. Since the example crucially utilises all of these peculiarities to exhibit the non-confluent behaviour, it is conceivable that imposing requirements to thwart just one of them may cause confluence to hold; we show in Section 4 that any proof of such a result must employ machinery significantly more powerful than that available to finitary and strongly convergent rewriting.

The fact that the set of rules of R does not have an upper bound on the depth of left-hand sides hints at a distinction from strongly convergent systems as already remarked in [5] where it is noted that the set of normal forms of such systems may differ in the Cauchy and strongly convergent settings. Another distinguishing factor arising from unbounded depth is that the analogue of Theorem 8 for Cauchy convergent rewriting will fail to hold, even for otherwise restricted systems, since hypercollapsingness is not preserved across Cauchy convergent rewriting with rules of unbounded depth:

Example 11 (*Hypercollapsingness is not preserved*). Consider the (non-confluent) orthogonal system with the infinite set of rules

$$\mathbf{f}\big(\mathbf{g}^k(\mathbf{c}), x\big) \to x \quad \text{for all } k \in \omega.$$

Starting from the term

$$s = \mathbf{f}\big(\mathbf{c}, \mathbf{f}\big(\mathbf{g}(\mathbf{c}), \mathbf{f}\big(\mathbf{g}\big(\mathbf{g}(\mathbf{c})\big), \ldots\big)\big)\big)$$

we have the convergent derivation

$$\underline{\mathbf{f}}\big(\mathbf{c}, \mathbf{f}\big(\mathbf{g}(\mathbf{c}), \mathbf{f}\big(\mathbf{g}\big(\mathbf{g}(\mathbf{c})\big), \ldots\big)\big)\big) \to \underline{\mathbf{f}}\big(\mathbf{g}(\mathbf{c}), \mathbf{f}\big(\mathbf{g}\big(\mathbf{g}(\mathbf{c})\big), \ldots\big)\big) \to \cdots$$

of length ω with limit

$$t = \mathbf{f}\big(\mathbf{g}^{\omega}, \mathbf{f}(\mathbf{g}^{\omega}, \ldots)\big)$$

which is a normal form, hence not hypercollapsing.

4. Residuals and the Transfinite Parallel Moves Lemma

The traditional stepping stone for proving confluence of orthogonal systems is the development of a suitable theory of *residuals*: the notion of "what happens" to redexes across contractions of other redexes. Definition of a residual relation is non-trivial in Cauchy convergent rewriting, as witnessed by the following well-known example:

Example 12. Consider the system

$$\mathbf{a} \to \mathbf{b},$$
$$\mathbf{f}(x, y) \to \mathbf{f}(y, x)$$

and the derivation of length ω:

$$\underline{\mathbf{f}}(\mathbf{a}, \mathbf{a}) \to \underline{\mathbf{f}}(\mathbf{a}, \mathbf{a}) \to \cdots$$

in which, though the derivation clearly converges to $\mathbf{f}(\mathbf{a}, \mathbf{a})$, it is not obvious what the residuals of the overlined $\mathbf{a}$ in $\mathbf{f}(\bar{\mathbf{a}}, \mathbf{a})$ should be in the limit.

The present section aims to show that the problems inherent in defining residuals in the presence of Cauchy convergence go beyond the vague notion that "subterm structure is lost" that one gets from the previous example. The essence of the problem can be seen from the example, however: the "commuting" behaviour of the arguments x and y of the rule $\mathbf{f}(x, y) \to \mathbf{f}(y, x)$ makes the rule act as if it were non-right-linear when derivations of infinite length are considered—as we shall see in a moment, the overlined $\mathbf{a}$ is in effect "duplicated" by the derivation.

In finitary and strongly convergent rewriting, the "Parallel Moves Lemma" for orthogonal systems states that for any two sets, $\mathcal{U}$ and $\mathcal{V}$, each composed of pairwise parallel redexes of the same term, the following *fundamental square* commutes:

$$\begin{tikzcd} \cdot \arrow[r, "\mathcal{U}"] \arrow[d, "\mathcal{V}"'] & \cdot \arrow[d, dotted, "\mathcal{V}/\mathcal{U}"] \\ \cdot \arrow[r, dotted, "\mathcal{U}/\mathcal{V}"'] & \cdot \end{tikzcd}$$

where all sides consist of complete developments of the indicated sets, and $\mathcal{U}/\mathcal{V}$ and $\mathcal{V}/\mathcal{U}$ can be chosen to consist of pairwise parallel redexes whenever $\mathcal{U}$ and $\mathcal{V}$ do. Confluence proofs in finitary rewriting often utilize the fundamental square to prove confluence by "tiling" a fork $t' \twoheadleftarrow s \twoheadrightarrow\!\!- t$, thus:

$$
\begin{tikzcd}
s \arrow[r] \arrow[d] & \cdot \arrow[r, dotted, no head] \arrow[d] & \cdot \arrow[r] \arrow[d] & \cdot \arrow[r, dotted, no head] \arrow[d] & t \arrow[d] \\
\cdot \arrow[r] \arrow[d, dotted, no head] & \cdot \arrow[r, dotted, no head] \arrow[d, dotted, no head] & \cdot \arrow[r] \arrow[d, dotted, no head] & \cdot \arrow[r, dotted, no head] \arrow[d, dotted, no head] & \cdot \arrow[d, dotted, no head] \\
\cdot \arrow[r] \arrow[d] & \cdot \arrow[r, dotted, no head] \arrow[d] & \cdot \arrow[r] \arrow[d] & \cdot \arrow[r, dotted, no head] \arrow[d] & \cdot \arrow[d] \\
\cdot \arrow[r] \arrow[d, dotted, no head] & \cdot \arrow[r, dotted, no head] \arrow[d, dotted, no head] & \cdot \arrow[r] \arrow[d, dotted, no head] & \cdot \arrow[r, dotted, no head] \arrow[d, dotted, no head] & \cdot \arrow[d, dotted, no head] \\
t' \arrow[r] & \cdot \arrow[r, dotted, no head] & \cdot \arrow[r] & \cdot \arrow[r, dotted, no head] & \cdot
\end{tikzcd}
$$

This is, essentially, also possible in strongly convergent transfinite rewriting since the upper horizontal and left vertical sides can be chosen to be of length $\leqslant \omega$ by the Compression Lemma. However, tiling by the fundamental square does not suffice for reasoning in the presence of Cauchy convergence: there is no Compression Lemma, and the topmost horizontal derivation thus cannot necessarily be chosen to have length $\leqslant \omega$; there will be limit ordinals "along the way", and tiling using the fundamental diagram will not in general be possible.

Thus, we want—at the very least—a "Transfinite" version of the Parallel Moves Lemma to hold:

$$
\begin{tikzcd}
s \arrow[r, "S"] \arrow[d, "\mathcal{U}"'] & t \arrow[d, dotted, "\mathcal{U}/S"] \\
s' \arrow[r, dotted, "S/\mathcal{U}"'] & t'
\end{tikzcd}
$$

in which S is any derivation. It is immediately obvious from the counterexample of Section 3 that this diagram in general *does not* commute, regardless of which definition of residual we choose (take $\mathcal{U}$ to be the singleton set consisting of the overlined redex in $\mathbf{f}(\mathbf{c}, \mathbf{a}, \bar{\mathbf{a}})$). However, it may at first appear conceivable that a usable Transfinite Parallel Moves Lemma could hold for a set of suitably restricted iTRSs.

The following examples will demonstrate that any notion of residual relation designed to make the Transfinite Parallel Moves Lemma hold will *either* have to part with a number of properties enjoyed in the finitary and strongly convergent cases, *or* fail to have the Transfinite Parallel Moves Lemma hold for some very simple systems. In each example we consider a derivation S and a set of redexes $\mathcal{U}$ such as the one in the diagram above, and argue that there can only be one choice for $\mathcal{U}/S$ if the diagram is to commute.

Example 13 (*Parallelness of redexes is not preserved across derivation*). Consider the orthogonal, non-collapsing, right-linear system

$$\mathbf{f}(x, y, z) \to \mathbf{f}\big(y, x, \mathbf{h}(z)\big),$$
$$\mathbf{a}(x) \to \mathbf{b}(x),$$
$$\mathbf{m}(x) \to \mathbf{n}(x)$$

and consider further the term $s = \mathbf{f}(\bar{\mathbf{a}}(\mathbf{m}(\mathbf{c})), \mathbf{a}(\overline{\mathbf{m}}(\mathbf{c})), \mathbf{c})$, the doubleton set $\mathcal{U}$ of overlined redexes in s (observe that the redexes in $\mathcal{U}$ are parallel), and the following derivation, S, of length ω:

$$
\begin{aligned}
s &\to \underline{\mathbf{f}}\big(\mathbf{a}\big(\mathbf{m}(\mathbf{c})\big), \mathbf{a}\big(\mathbf{m}(\mathbf{c})\big), \mathbf{h}(\mathbf{c})\big) \\
&\to \underline{\mathbf{f}}\big(\mathbf{a}\big(\mathbf{m}(\mathbf{c})\big), \mathbf{a}\big(\mathbf{m}(\mathbf{c})\big), \mathbf{h}^2(\mathbf{c})\big) \\
&\to \cdots.
\end{aligned}
$$

The limit of which is $\mathbf{f}(\mathbf{a}(\mathbf{m}(\mathbf{c})), \mathbf{a}(\mathbf{m}(\mathbf{c})), \mathbf{h}^\omega)$. Contracting the redexes of $\mathcal{U}$ in s yields the term $\mathbf{f}(\mathbf{b}(\mathbf{m}(\mathbf{c})), \mathbf{a}(\mathbf{n}(\mathbf{c})), \mathbf{c})$. It is clear that the only possible common derivative of

$$\mathbf{f}\big(\mathbf{a}\big(\mathbf{m}(\mathbf{c})\big), \mathbf{a}\big(\mathbf{m}(\mathbf{c})\big), \mathbf{h}^\omega\big)$$

and

$$\mathbf{f}\big(\mathbf{b}\big(\mathbf{m}(\mathbf{c})\big), \mathbf{a}\big(\mathbf{n}(\mathbf{c})\big), \mathbf{c}\big)$$

is

$$\mathbf{f}\big(\mathbf{b}\big(\mathbf{n}(\mathbf{c})\big), \mathbf{b}\big(\mathbf{n}(\mathbf{c})\big), \mathbf{h}^\omega\big).$$

Thus, $\mathcal{U}/S$ must consist of the overlined redexes in $\mathbf{f}(\bar{\mathbf{a}}(\overline{\mathbf{m}}(\mathbf{c})), \bar{\mathbf{a}}(\overline{\mathbf{m}}(c)), \mathbf{h}^\omega)$ which clearly are not pairwise parallel.

Example 14 (*Duplication may occur in right-linear systems*). In the orthogonal, non-collapsing, right-linear system

$$\mathbf{a} \to \mathbf{b},$$
$$\mathbf{f}(x, y, z) \to \mathbf{f}\big(y, x, \mathbf{g}(z)\big)$$

the following derivation, S, of length ω:

$$\underline{\mathbf{f}}(\mathbf{a}, \mathbf{a}, \mathbf{c}) \to \underline{\mathbf{f}}\big(\mathbf{a}, \mathbf{a}, \mathbf{g}(\mathbf{c})\big) \to \underline{\mathbf{f}}\big(\mathbf{a}, \mathbf{a}, \mathbf{g}\big(\mathbf{g}(\mathbf{c})\big)\big) \to \cdots$$

converges to $\mathbf{f}(\mathbf{a}, \mathbf{a}, \mathbf{g}^\omega)$, whereas contraction of the over lined redex in $\mathbf{f}(\bar{\mathbf{a}}, \mathbf{a}, \mathbf{c})$ yields $\mathbf{f}(\mathbf{b}, \mathbf{a}, \mathbf{c})$. The only possible common derivative of these two terms is $\mathbf{f}(\mathbf{b}, \mathbf{b}, \mathbf{g}^{\omega})$. Hence, for the Transfinite Parallel Moves Lemma to hold, *both* occurrences of a in $\mathbf{f}(\mathbf{a}, \mathbf{a}, \mathbf{g}^{\omega})$ must be counted as residuals of the over lined redex in $\mathbf{f}(\bar{\mathbf{a}}, \mathbf{a}, \mathbf{c})$.

Example 15 (*Redexes may have more than one ancestor*). Consider the system of the previous example and the term $\mathbf{f}(\bar{\mathbf{a}}, \bar{\mathbf{a}}, \mathbf{c})$ with the redexes of the doubleton set $\mathcal{U}$ overlined. Performing ω outermost contractions yields the term $\mathbf{f}(\mathbf{a}, \mathbf{a}, \mathbf{g}^{\omega})$, whereas contracting the set $\mathcal{U}$ yields $\mathbf{f}(\mathbf{b}, \mathbf{b}, \mathbf{c})$. By the reasoning of the last example, *both* occurrences of $\mathbf{a}$ in $\mathbf{f}(\mathbf{a}, \mathbf{a}, \mathbf{g}^{\omega})$ must be residuals of *each* of the overlined redexes in the starting term.

On a concluding note, we turn our attention to the derivation S in the diagram of the Transfinite Parallel Moves Lemma. In finitary and strongly convergent rewriting, the derivation $S/\mathcal{U}$ can be taken as a *projection* of S, i.e., essentially "the same" rewrite steps as S, taking into account that $\mathcal{U}$ has already been developed. It is an easily obtained fact that we are not as fortunate when considering Cauchy convergence: Close inspection of the proof of Proposition 4, where we sought a common derivative of $\mathbf{f}(\mathbf{b}, \mathbf{a}, \mathbf{c})$ and $\mathbf{f}(\mathbf{a}, \mathbf{a}, \mathbf{g}^{\omega})$, reveals that the standard projection of the considered derivation S across $\mathbf{f}(\underline{\mathbf{a}}, \mathbf{a}, \mathbf{c}) \to \mathbf{f}(\mathbf{b}, \mathbf{a}, \mathbf{c})$ is not convergent—we must first perform the rewrite step $\mathbf{f}(\mathbf{b}, \underline{\mathbf{a}}, \mathbf{c}) \to \mathbf{f}(\mathbf{b}, \mathbf{b}, \mathbf{c})$ to obtain a convergent derivation. Variations on the example can be constructed that require us to perform an infinite number of such "corrective" rewrite steps before performing the rewrite steps of the projection, or to permute the steps of S. It is evident that any workable definition of residual must be accompanied by a suitable—and new—notion of projection that takes the above into account.

References

[1] F. Baader, T. Nipkow, Term Rewriting and All That, Cambridge Univ. Press, Cambridge, UK, 1998.

[2] N. Dershowitz, J.-P. Jouannaud, Rewrite systems, in: J. van Leeuwen (Ed.), Handbook of Theoretical Computer Science, vol. B, Elsevier, Amsterdam, 1990, pp. 243–320.

[3] N. Dershowitz, S. Kaplan, D. Plaisted, Rewrite, rewrite, rewrite, rewrite, rewrite ..., Theoret. Comput. Sci. 83 (1) (1991) 71–96.

[4] R. Kennaway, F.-J. de Vries, Infinitary rewriting, in: Term Rewriting Systems, Cambridge Tracts Theoret. Comput. Sci., vol. 55, Cambridge Univ. Press, Cambridge, UK, 2003, pp. 668–711, Chapter 12.

[5] R. Kennaway, J.W. Klop, R. Sleep, F.J. de Vries, Transfinite reductions in orthogonal term rewriting, Inform. and Comput. 119 (1) (1995) 18–38.

[6] R. Kennaway, V. van Oostrom, F.J. de Vries, Meaningless terms in rewriting, J. Funct. Logic Programming 1 (1999) 1–34.

[7] J.W. Klop, Term rewriting systems, in: S. Abramsky, D. Gabbay, T. Maibaum (Eds.), Handbook of Logic in Computer Science, vol. 2, Oxford Univ. Press, London, 1992, pp. 1–116.

[8] E. Mendelson, Introduction to Mathematical Logic, fourth ed., Chapman and Hall, London/New York, 1997.

[9] TeReSe, Term Rewriting Systems, Cambridge Tracts Theoret. Comput. Sci., vol. 55, Cambridge Univ. Press, Cambridge, UK, 2003.