Centrum voor Wiskunde en Informatica

REPORT*RAPPORT*

SEN

Software Engineering

Software ENgineering

Vicious circles in rewriting systems

J. Ketema, J.W. Klop, V. van Oostrom

REPORT SEN-E0427 DECEMBER 2004

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).
CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

ISSN 1386-369X

Vicious circles in rewriting systems

ABSTRACT
We continue our study of the difference between Weak Normalisation (WN) and Strong Normalisation (SN). We extend our earlier result that orthogonal TRSs with the property WN do not admit cyclic reductions, into three distinct directions: (i) to the higher-order case, where terms may contain bound variables, (ii) to the weakly orthogonal case, where rules may have (trivial) conflicts, and (iii) to weak head normalisation (WHN), where terms have head normal forms. By adapting the techniques introduced for each of the three extensions separately, we even are able to show the result generalises to each pair of combinations and to various lambda-calculi. The combination of all three extensions remains open however.

2000 Mathematics Subject Classification: 68Q42
Keywords and Phrases: Term rewriting systems;cyclic reductions;normalisation;functional programming

Vicious Circles in Rewriting Systems

Jeroen Ketema[1]

Department of Theoretical Computer Science, Vrije Universiteit, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands

Jan Willem Klop[2]

Department of Theoretical Computer Science, Vrije Universiteit, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands; CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands; Department of Computer Science, Radboud Universiteit Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

Vincent van Oostrom[3]

Department of Philosophy, Universiteit Utrecht, Heidelberglaan 8, 3584 CS Utrecht, The Netherlands

---

Abstract

We continue our study of the difference between Weak Normalisation (WN) and Strong Normalisation (SN). We extend our earlier result that orthogonal TRSs with the property WN do not admit cyclic reductions, into three distinct directions:
(i) to the higher-order case, where terms may contain bound variables,
(ii) to the weakly orthogonal case, where rules may have (trivial) conflicts, and
(iii) to weak head normalisation (WHN), where terms have head normal forms.
By adapting the techniques introduced for each of the three extensions separately, we even are able to show the result generalises to each *pair* of combinations and to various λ-calculi. The combination of all *three* extensions remains open however.

Key words: Rewriting systems, cyclic reductions, normalisation.

---

[1] Email: `jketema@cs.vu.nl`
[2] Email: `jwk@cs.vu.nl`
[3] Email: `oostrom@phil.uu.nl`

1 Introduction

.

We continue our study of the difference between Weak Normalisation (WN) and Strong Normalisation (SN) started in [10], to which we refer the reader for ample motivation. There we showed that although WN in general does not imply the absence of *infinite* reductions, it does imply *acyclicity* (AC), i.e. the absence of *cyclic* reductions in the case of orthogonal first-order term rewriting systems (TRS). The prototypical such TRS witnessing both phenomena is:

$$\begin{aligned} a &\to f(a) \\ f(x) &\to b \end{aligned}$$

with typical reduction ladder (see [10]) as in Figure 1. As exemplified by the

$$\begin{tikzcd}
a \arrow[r] & f(a) \arrow[r] \arrow[d] & f(f(a)) \arrow[r] \arrow[d] & f(f(f(a))) \arrow[r] \arrow[d] & f(f(f(f(a)))) \arrow[r] \arrow[d] & {} \\
 & b & f(b) \arrow[l] & f(f(b)) \arrow[l] & f(f(f(b))) \arrow[l] & {} \arrow[l]
\end{tikzcd}$$

Figure 1. Infinite reduction ladder in orthogonal TRS which is WN but not SN

figure, this TRS is WN. By the main result of [10], it therefore does not admit cyclic reductions, although it obviously does admit infinite reductions. From acyclicity we conclude, using that there are only finitely many terms of a given size, that the terms along such infinite reductions must grow unboundedly in size, as witnessed by the reduction ladder in the figure.

In this paper, we extend the WN $\Rightarrow$ AC result, as obtained in [10] for orthogonal first-order term rewriting systems, into three distinct directions:

(i) to the higher-order case, where terms may contain bound variables,

(ii) to the weakly orthogonal case, where rules may have (trivial) conflicts, and

(iii) to weak head normalisation (WHN), where terms have head normal forms.

In particular, the first of these extensions pertains to the $\lambda\beta$-calculus [4]. In itself this is not interesting as the $\lambda\beta$-calculus is not WN. We will show however that our result carries over to its sub-calculi and in particular to typed λ-calculi, for which WN often *can* be shown. Therefore, we provide a *partial* answer to:

Conjecture 1.1 (Barendregt–Geuvers–Klop) *WN $\Rightarrow$ SN, for typed $\lambda\beta$.*

In particular, although we do *not* show that infinite reductions do not exist in such a WN typed $\lambda\beta$-calculus, it *does* follow that if an infinite reduction would exist, then the situation is as in the prototypical TRS above: the infinite reduction is acyclic and the terms along it grow unboundedly in size.

We also show that the result pertains to each *pair* of combinations of the individual extensions presented above. In particular, combining the first two extensions, the result pertains to weakly orthogonal fully-extended higher-order pattern rewrite systems (PRSs [13]). Although the $\lambda\beta\eta$-calculus and its typed sub-calculi are the prototypical examples of weakly orthogonal PRSs, they are *not* fully-extended [11], so our result does not pertain immediately to them. However, we show that our methods can be adapted to these as well, yielding the same partial answer to the Barendregt–Geuvers–Klop conjecture for typed $\lambda\beta\eta$-calculi as for such calculi without η.

The combination of all three extensions remains open however:

Conjecture 1.2 *WHN $\Rightarrow$ AC, for weakly orthogonal PRSs.*

We conclude this introduction with an overview of the rest of the paper. In Section 2 we recapitulate the part of the theory of projections and residuals for first-order orthogonal TRSs [17, Ch. 8] pertinent to this paper, and adapt it to the higher-order and the weakly orthogonal case. The subsequent Sections 3–5 constitute the technical core of the paper. In particular, in Section 3 we adapt various proof techniques from the literature, see e.g. [10,16,12], to obtain each of the three individual generalisations mentioned above. In Section 4, we then consider the (three) pair-wise combinations of each of the separate generalisations. Of these, only the combination of weak orthogonality with weak head normalisation turns out to be non-trivial. The proof of this combination, for first-order term rewriting systems, will take up the largest part of the paper. The proof introduces some novel techniques based on the projection of cyclic reductions and on covering clusters of overlapping redex-patterns by chains of such. Finally, in Section 5, we turn our attention to λ-calculi, in particular to $\lambda\beta$-calculi, $\lambda\beta\eta$-calculi, and the λ-calculi with explicit substitutions $\lambda\mathtt{x}^-$ and $\lambda\sigma$. Sections 3–5 can be read independently from each other.

Throughout, the rewriting systems considered are left-linear.

2 Preliminaries on projection

In this section we present the kind of rewriting systems considered in this paper, and show that they have projection (in a technical sense to be defined). Projections will be the main tool employed in the subsequent sections. We illustrate projections for each kind of rewriting system by means of examples, which will serve as running examples throughout the rest of the paper.

This section is mainly a recapitulation of known results, with the exception of the particular projection employed in the weakly orthogonal case.

2.1 ARSs and projections

We briefly recapitulate the notions of cycle, projection and extension, which are crucial in this paper, at the level of abstract rewriting systems (ARSs). In particular, we first define projections of steps, which will then be generalised to projections of reductions. See [17,3] for introductions to ARSs. We employ ARSs as defined in [17, Def.8.2.2], a, b, c, … to range over objects, and ϕ, ψ, χ, … to range over steps. We employ σ, τ, υ, … to denote (finite or infinite) reductions.

A *loop* is a step having the same source and target. A non-empty reduction in a given ARS $\to$ is a *cycle*, if it is a loop in the associated reduction-ARS $\twoheadrightarrow$. If the source or target of a loop (cycle) is a, we will speak of a loop (cycle) *on* a. In this paper, we will be interested in *acyclicity* (AC) of rewriting systems, i.e. in the absence of cycles.

Recall, e.g. from [17, Def. 1.1.8(v)], that an ARS has the *diamond property* if any pair of co-initial steps can be completed into a 'diamond' by a pair of co-final steps (see Figure 2 left). We say that an ARS has projection if there

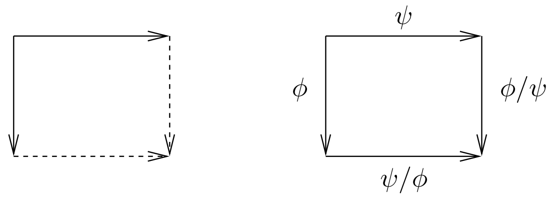

Figure 2. Diamond property and projection

is a *functional* way for completion.

Definition 2.1 An ARS $\to$ is said to *have projection* if there is an operation / from pairs of co-initial steps ϕ, ψ to steps such that (see Figure 5 right):

- the target of ψ is the source of ϕ/ψ, and
- the targets of ϕ/ψ and ψ/ϕ coincide.

ϕ/ψ is the *residual of* ϕ *after* ψ, which is obtained by *projecting* ψ *over* ϕ.

An ARS has projection if and only if it has the diamond property.

Proposition 2.2 *If* $\to$ *has the diamond projection, then it has projection.*

Proof In order to define the function /, we consider an arbitrary *unordered* pair ϕ, ψ of co-initial $\to$-steps. Ordering the pair arbitrarily, say into ϕ, ψ, the resulting *peak* can be completed into a diamond by some *valley*. Choose among these valleys an arbitrary one, say the *ordered* pair ψ', ϕ' of co-final $\to$-steps. Then in case ϕ is distinct from ψ, we can simply set $\phi/\psi = \phi'$ and $\psi/\phi = \psi'$. In case ϕ and ψ are identical we must choose, in order to remain functional, arbitrarily only one of these, say $\phi/\phi = \phi'$. $\square$

In particular, a constructive proof that an ARS has the diamond property yields a projection.

Remark 2.3 Note that by the diamond property, if there is some step ϕ *to* an object a, then there is some step *from* that object a as well. This follows, since completing the peak ϕ, ϕ, yields at least one such step. Hence the normal forms are the isolated points of the reduction graph of such an ARS!

Remark 2.4 Abstract residual systems as introduced in [17, Sec. 8.7], can be seen as ARSs having projection such that the projection operation moreover satisfies some algebraic laws. We expect that of these laws the so-called identity laws hold in all our applications below, after enriching ARSs with looping steps, but conjecture that the so-called cube law cannot be made to hold in general, cf. Remark 2.26. Hence we will not base our theory on residual systems. (However, orthogonal term rewriting systems *do* yield residual systems.)

Having projection is inherited by certain sub-ARSs [17, Def. 8.2.5(i)].

Lemma 2.5 *Let $\rightsquigarrow$ be a sub-ARS of the ARS $\to$ which has projection $/$.*

- *$\rightsquigarrow$ has projection, if it is closed under $/$-projections.*
- *$\rightsquigarrow$ is closed under $/$-projections, if it is closed under $\twoheadrightarrow$-reduction.*

Hence, given a subset of the objects of an ARS having projections, to show the sub-ARS induced by the subset has projections, it suffices to show that it is closed under reductions.

Projection extends from steps to reductions by *tiling* (see Figure 3).

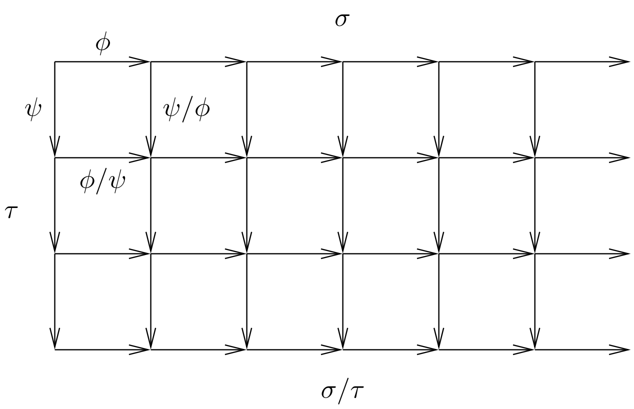

Figure 3. Projection of infinite reduction σ over finite reduction τ by tiling

Definition 2.6 Let $\to$ be an ARS having projection $/$. Then the projection of a finite or infinite reduction over a finite one is defined by (cf. [9]):

- $\sigma/\emptyset = \sigma$,
- $\sigma/(\psi \cdot \tau) = (\sigma/\psi)/\tau$,
- $\emptyset/\psi = \emptyset$, and
- $(\phi \cdot \sigma)/\psi = (\phi/\psi) \cdot (\sigma/(\psi/\phi))$.

In this paper, we will be interested in projections of certain infinite reductions and cycles, such that their residuals have the same shape again.

Definition 2.7 A property of reductions is said to be *preserved* by projection, if whenever it holds for a given finite or infinite reduction σ, then it also holds for the residual σ/τ of σ after any finite co-initial reduction τ.

Note that the property of being of given length is preserved by projection.

In Section 4 the concept of a reduction being an extension of another, co-initial one, will play an important role. Since it seems to have some independent interest (cf. co-finality [17, Def. 1.1.15(iv)]), we present it here.

Definition 2.8 Let ϕ, ψ be co-initial steps in some ARS. If there is a reduction from the target of ϕ to the target of ψ, then ψ is an *extension* of ϕ. We say that the ARS *creates* a given property, if for every step ϕ, there exists an extension of ϕ having the property.

It is easy to see that extension is reflexive and transitive, i.e. a quasi-order. Note that, starting with the ARS $\twoheadrightarrow$ of *reductions*, a reduction σ trivially extends any of its prefixes, including σ itself and the empty reduction from its source. Moreover, to show that τ is an extension of σ, it suffices to show there is a reduction from the target of σ to *some* object on τ.

Lemma 2.9 *Let $\to$ be an ARS having projection. Then if $\to$ creates a given property, then the reduction-ARS $\twoheadrightarrow$ creates the property (extended stepwise) as well.*

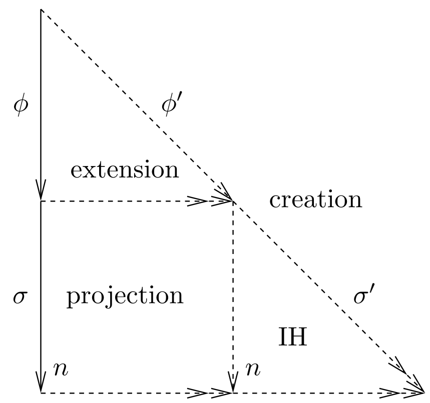

Figure 4. Proof of Lemma 2.9

Proof The proof is by induction on the length of a reduction. If the reduction is empty, then there is nothing to prove. Otherwise, let the reduction consist of a step ϕ followed by a reduction σ say of length n (see Figure 4). By assumption there are a step ϕ' having the property, and a reduction τ from the target of ϕ to the target of ϕ'. By the ARS having projection, there are reductions τ/σ and σ/τ completing σ, τ into a diamond. As noted in the previous paragraph σ/τ has length n again, so we may apply the induction hypothesis to it, yielding a reduction σ' extending σ/τ via reduction τ' having the property. Therefore, all steps of the reduction consisting of ϕ' followed by σ' have the property, and this reduction is an extension via the reduction consisting of τ/σ and τ'. $\square$

Note that the lemma generalises to infinite reductions (not needed here). In this paper, we will be interested in creation of properties for computations of results, i.e. for reductions whose target is an element of a set of *results*, e.g. the set of head normal forms. As sets of results are closed under reduction, extensions of such reductions will be computations of results again.

2.2 Orthogonal TRSs and projections

As it will be the basis for our generalisations, we briefly recapitulate the well-known fact that both parallel step and multi-steps in orthogonal TRSs have projection [9,12,17].

We assume basic knowledge of the theory of first-order term rewriting systems (TRSs) see [17,3]. We employ TRSs as defined in [17, Def. 8.2.15].

Example 2.10 Consider the term $f(a)$ in the orthogonal TRS with rules:

$$\begin{aligned} a &\to b \\ f(x) &\to g(x) \end{aligned}$$

There are two steps having $f(a)$ as source, obtained by contracting either the a-redex-pattern or the f-redex-pattern:

$$\begin{aligned} &\phi{:}f(\underline{a}) \to f(b) \\ &\psi{:}\underline{f}(a) \to g(a) \end{aligned}$$

The targets $f(b)$ and $g(a)$ of the steps are distinct. Still they do have a common reduct $g(b)$, obtained by contracting the 'other' redex-patterns, i.e. the f-redex-pattern and the a-redex-pattern, respectively (see Figure 5 left):

$$\begin{aligned} &\psi/\phi{:}\underline{f}(b) \to g(b) \\ &\phi/\psi{:}g(\underline{a}) \to g(b) \end{aligned}$$

Intuitively, the f-redex-pattern in $f(b)$ can be thought of as the *residual of* the f-redex-pattern in the original term $f(a)$, *after* the step ϕ. Similarly, we have that the a-redex-pattern in $g(a)$ is the residual of the a-redex-pattern in $f(a)$ after the step ψ.

The phenomenon witnessed by the example is that of projection as defined in Definition 2.1: for a given peak $s \leftarrow_{\phi} t \to_{\psi} u$ from some term t, a common reduct v is found by *projecting* each step over the other and performing the respective residual steps $s \to_{\psi/\phi} v \leftarrow_{\phi/\psi} u$. Because of Remark 2.3, the single step ARS $\to_{\mathcal{R}}$ [17, Prop. 8.2.22(i)] associated to a TRS $\mathcal{R}$ in practice never has projection. However its associated multi-step and parallel step ARSs $\multimap_{\mathcal{R}}$ and $\mathrel{+\!\!\!\!\!\!\to}_{\mathcal{R}}$ [17, Fig. 8.8] *do* have projection, in case $\mathcal{R}$ is orthogonal. [4]

[4] Therefore, neither the multi-step nor the parallel step ARS has normal forms. For our

$$
\begin{tikzcd}
f(a) \arrow[r, "\psi"] \arrow[d, "\phi"'] & g(a) \arrow[d] \\
f(b) \arrow[r] & g(b)
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
I(Kz) \arrow[r, "\psi"] \arrow[d, "\phi"'] & Kz \arrow[d] \\
I(\lambda y.z) \arrow[r] & \lambda y.z
\end{tikzcd}
$$

Figure 5. Projection for a TRS and for the $\lambda\beta$-calculus

Example 2.11 Consider the co-initial parallel steps (which, incidentally, are single steps as well) $\phi{:}f(\underline{a}) \mathrel{\|\!\!\!\longrightarrow} f(b)$ and $\psi{:}\underline{f}(a) \mathrel{\|\!\!\!\longrightarrow} g(a,a)$ in the orthogonal TRS with rules

$$
\begin{aligned}
a &\to b \\
f(x) &\to g(x,x)
\end{aligned}
$$

First, note that the projection of ϕ over itself, yields the looping parallel step $\phi/\phi{:}b \mathrel{\|\!\!\!\longrightarrow} b$, contracting 'zero redexes in parallel'.

Next, note that projecting ϕ and ψ over each other indeed yields parallel steps $\phi/\psi{:}\underline{f}(b) \mathrel{\|\!\!\!\longrightarrow} g(b,b)$ and $\psi/\phi{:}g(\underline{a},\underline{a}) \mathrel{\|\!\!\!\longrightarrow} g(b,b)$ having the same target.

Viewing ϕ and ψ as multi-steps (recall that parallel steps are a special case of multi-steps) yields exactly the same construction. However, note that although there does not exist a *parallel* step from the initial term $f(a)$ to the final term $g(b,b)$, there does exist such a multi-step, namely the step $\phi \cup \psi{:}\underline{f}(\underline{a}) \mathrel{\circ\!\!\!\longrightarrow} g(b,b)$ contracting the *union* of the redexes contracted in ϕ,ψ.

Lemma 2.12 *Let $\mathcal{R}$ be an orthogonal TRS. Then*

- *the multi-step ARS $\mathrel{\circ\!\!\!\longrightarrow}_{\mathcal{R}}$ has projection,*
- *the parallel step ARS $\mathrel{\|\!\!\!\longrightarrow}_{\mathcal{R}}$ has projection, and*
- *the projection of a parallel step over a multi-step is a parallel step.*

Proof For the first two items, see [17, Prop. 8.7.7] where the stronger property is shown that these ARSs constitute so-called residual systems. For the third item, employ that any multi-step can be developed into a series of parallel steps and (repeatedly) use the second item. $\square$

A parallel step or multi-step contracting an empty set of redex-patterns is said to be *empty*. If the residual of the reduction σ after τ consists only of empty steps, then τ is said to *eliminate* or *absorp* σ.

2.3 Orthogonal PRSs and projection

We show that orthogonal higher-order pattern rewriting systems (PRSs) [13] have projection. Here, we provide only some intuitions for higher-order rewriting and examples illustrating PRSs. We refer the reader to [17, Chapter 11], in particular to Definition 11.2.23, for a gentle introduction and concrete definitions.

purposes, and more generally for proving confluence results, this is irrelevant.

Higher-order pattern rewriting systems are *higher-order* in the sense that the objects being rewritten are terms which may contain bound variables. More precisely, the terms are simply (Church) typed λ-terms modulo $\alpha\beta\eta$ over some (user-defined) alphabet of simply typed constants. Rewrite rules consist of pairs of terms as usual, but for first-order rewrite theory to carry over, left-hand sides are restricted to so-called *rule-patterns*, making that critical pairs and hence orthogonality can be determined.

Example 2.13 The terms of the untyped λ-calculus can be encoded as simply typed terms of type Trm, by including the simply typed constants

$\mathsf{app}{:}\mathsf{Trm}{\to}(\mathsf{Trm}{\to}\mathsf{Trm})$, and

$\mathsf{abs}{:}(\mathsf{Trm}{\to}\mathsf{Trm}){\to}\mathsf{Trm}$

into the alphabet and restricting variables to type Trm. To denote terms, we apply the usual notational conventions for the λ-calculus except that we *suppress* writing λs. For instance, the ordinary λ-term $(\lambda x.x)z$ is encoded as $\mathsf{app}(\mathsf{abs}\, x{:}\mathsf{Trm}.x)z$ of type Trm. Using this encoding, the ordinary *rule scheme* for β-reduction:

$$(\lambda x.M)N \to M[x{:=}N]$$

with M, N meta-variables is now rendered as a proper *rule*:

$$M{:}\mathsf{Trm}{\to}\mathsf{Trm}.N{:}\mathsf{Trm}.\mathsf{app}(\mathsf{abs}\, x{:}\mathsf{Trm}.Mx)N \;\to\; M{:}\mathsf{Trm}{\to}\mathsf{Trm}.N{:}\mathsf{Trm}.MN$$

Since the left-hand side is a rule-pattern [17, Def. 11.2.18], this is a PRS rule. (Note how substitution in the right-hand side of the original rule scheme is rendered by means of the application MN in the rule of the PRS.)

To see how λ-terms in this encoding are rewritten the β-rule, reconsider the λ-term $(\lambda x.x)z$ which can be rewritten using the β-rule scheme as:

$$(\lambda x.x)z \to z$$

For the PRS encoding we have that $\mathsf{app}(\mathsf{abs}\, x{:}\mathsf{Trm}.x)z$ expands to the term $(M{:}\mathsf{Trm}{\to}\mathsf{Trm}.N{:}\mathsf{Trm}.\mathsf{app}(\mathsf{abs}\, x{:}\mathsf{Trm}.Mx)N)(x{:}\mathsf{Trm}.x)z$ where the left-hand of the rule is explicitly present. Replacing in this term the left-hand side of the rule by its right-hand side yields $(M{:}\mathsf{Trm}{\to}\mathsf{Trm}.N{:}\mathsf{Trm}.MN)(x{:}\mathsf{Trm}.x)z$, which reduces to z. Combining the expansion with the subsequent replacement and reduction yields the expected PRS step:

$$\mathsf{app}(\mathsf{abs}\, x{:}\mathsf{Trm}.x)z \to z$$

For the sake of readability, we will henceforth employ ordinary notations for λ-terms, where in fact we mean their higher-order encodings.

Orthogonality of PRSs is defined, as in the first-order case, as the absence of so-called critical pairs [17, Def. 11.6.17]. Therefore projections can be computed by tracing residuals, as in the first-order case,

Example 2.14 The encoding of the $\lambda\beta$-calculus as a PRS is orthogonal since it has no critical pairs. More specifically, the left-hand side of the β-rule does not overlap itself, it is orthogonal, so we conclude that the ordinary $\lambda\beta$-calculus 'is' an orthogonal PRS.

Consider the term $I(Kz) = (\lambda x.x)((\lambda xy.x)z)$ in the $\lambda\beta$-calculus. There are two β-reduction steps having $I(Kz)$ as source, obtained by contracting either the K-redex or the I-redex:

$$\phi{:}I(\underline{K}z) \to_\beta I(\lambda y.z)$$
$$\psi{:}\underline{I}(Kz) \to_\beta Kz$$

The targets $I(\lambda y.z)$ and Kz of the steps are distinct. Still they do have a common reduct $\lambda y.z$, obtained by contracting the 'other' redexes, i.e. the I-redex and the K-redex, respectively (see Figure 5 right):

$$\psi/\phi{:}\underline{I}(\lambda y.z) \to_\beta \lambda y.z$$
$$\phi/\psi{:}\underline{K}z \to_\beta \lambda y.z$$

Intuitively, the I-redex in $I(\lambda y.z)$ can be thought of as the *residual of* the I-redex in the original term $I(Kz)$, *after* the step ϕ. Similarly, the K-redex in Kz is the residual of the K-redex in the original term $I(Kz)$ after ψ.

Lemma 2.15 *$\multimap\!\!\to_{\mathcal{H}}$ has projection, if $\mathcal{H}$ is an orthogonal PRS.*

Proof See [17, Sec. 11.6.2] □

Remark 2.16 Like for TRSs, the multi-step ARS of an orthogonal PRS even constitutes a residual system. However, unlike for TRSs, the associated parallel step ARS does *not* have projection in general for orthogonal PRSs. See [17, Exs. 11.6.20].

In particular the (PRS encoding of the) $\lambda\beta$-calculus has projection; a fact which is of course well-known as it is at the basis of the Tait–Martin-Löf method for proving confluence [4].

Often the encoding of a rewriting system as a PRS yields spurious terms, i.e. terms which are not the image of any object in the original system. For instance, the usual encoding of the $\lambda\beta$-calculus as a PRS allows for terms containing variables having arbitrary simple types (constructed from Trm), which are not the image of any ordinary λ-term. It is easy to see that a 1–1 correspondence is obtained when variables in terms are restricted to the base type Trm. As the resulting subset of higher-order terms is closed under reduction, the sub-ARS of $\multimap\!\!\to$ induced by it, has projection again by Lemma 2.5.

2.4 Weakly orthogonal PRSs and projection

A rewriting system is defined to be *weakly* orthogonal in case all its critical pairs are trivial. That is, left-hand sides are allowed to have overlap, but in case one redex overlaps another one, the result of contracting either of them is the same. Again for concreteness sake, we take [17, Def. 11.6.17(ii)] as definition of weak orthogonality.

Example 2.17 • The TRS with rules

$P(S(x)) \to x$

$S(P(x)) \to x$

is weakly orthogonal since both its two critical peaks are trivial:

$P(x) \leftarrow \underline{P(\overline{S}}\overline{(P}(x)) \to P(x)$

$S(x) \leftarrow \underline{S(\overline{P}}\overline{(S}(x)) \to S(x)$

- The PRS encoding of the $\lambda\beta\eta$-calculus is obtained from $\lambda\beta$-calculus by translating the ordinary η-rule scheme:

  $\lambda x.Mx \to M$, if $x \notin M$

  in the same way as the β-rule scheme before, yielding: [5]

  $M{:}\mathsf{Trm}.\mathsf{abs}\,x{:}\mathsf{Trm}.\mathsf{app}Mx \to M{:}\mathsf{Trm}.M$

  This resulting $\lambda\beta\eta$-calculus is weakly orthogonal since it has two critical peaks both of which are trivial. They are, reverting to standard notation:

  $MN \leftarrow (\lambda x.Mx)N \to MN$

  $\lambda y.M \leftarrow \lambda x.(\lambda y.M)x \to \lambda x.M[y{:=}x]$

  where $x \notin M$ in both peaks. (To see triviality of the second peak, recall that we consider terms modulo $\alpha\beta\eta$.)

In order to project multi-steps over one another in weakly orthogonal PRSs, we must decide how to deal with the possibility that one of the multi-steps contains a redex having overlap with a redex in the other multi-step.

Definition 2.18 Let t be a term in a weakly orthogonal PRS. The *maximal* multi-step t^* from t contracts the set of redex-patterns selected by the following procedure (see Figure 6): [6]

- Start with the empty set.
- Given a set Φ of selected redex-patterns, select a redex-pattern ϕ in t which does not have overlap with any redex-pattern in Φ and whose head is at maximal depth among those.
- If selection fails, return Φ. Otherwise, adjoin the selected step ϕ to the set Φ of selected redexes and repeat the previous step.

Note that this procedure is not necessarily deterministic since there may be distinct redex-patterns the heads of which have overlap (this is the case for the two redexes on the middle right in t in Figure 6). Putting a well-order on

[5] Because the object-variable M has type Trm, it *cannot* be instantiated by a term containing the displayed bound variable x. In this way the external side-condition that x may not occur in M of the ordinary η-rule, is internalised in its PRS encoding.

[6] In orthogonal rewriting theory in general, and for the $\lambda\beta$-calculus in particular, a maximal multi-step is also known as a *complete development* or a *Gross–Knuth* step. In first-order term rewriting it is also called a *full substitution* step. See [17].

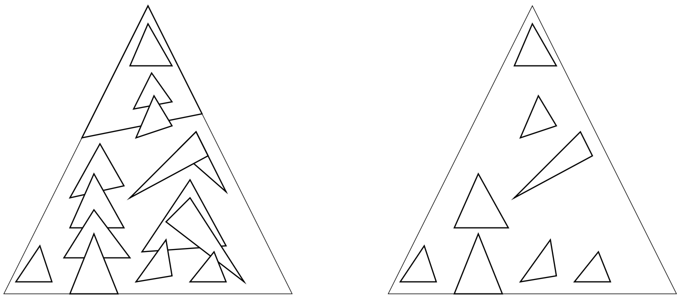

Figure 6. Term t with redex-patterns having overlap, and its maximal selection t^*.

the rules, and choosing a minimal one, is an easy way to make the procedure deterministic in such cases.

Example 2.19 We have drawn overlaps on the left in Figure 6 such that an inner redex-pattern is always on top of (*over*laps) an outer one. Maximisation then consists in repeatedly selecting all redex-patterns which are not covered, and removing with them all redex-patterns they (partially) cover, yielding the selection on the right in the figure.

Example 2.20 Consider the term $t = \overline{P(\underline{S(}}\underline{\overline{P(}}\overline{\underline{S(}}\underline{P(}x)))))$ in the TRS of Example 2.17. This term has the four redex-patterns indicated, which we number 1, 2, 3, and 4 from left to right. Executing the selection procedure then proceeds as follows:

- Start with the empty set $\emptyset$ of selected redexes.
- The set of redex-patterns not having overlap with $\emptyset$ is $\{1, 2, 3, 4\}$. Since 4 is at maximal depth among those, it is selected and adjoined to $\emptyset$.
- The set of redex-patterns not having overlap with $\{4\}$ is $\{1, 2\}$. Since 2 is at maximal depth among those, it is selected and adjoined to $\{4\}$.
- Since there is no redex-pattern which does not have overlap with $\{2, 4\}$, the latter set is returned.

Hence the maximal multi-step from t is $t^*{:}P(\underline{S(P(}\underline{S(P(}x))))) \mathrel{-\!\!\circ\!\!\to} P(x)$.

By the selection procedure, we immediately have that for *any* redex-pattern ϕ in t we may select some redex-pattern $\iota(\phi)$ in t^* the head of which has overlap with ϕ. Assuming that also this selection is deterministic, for a given ϕ, this yields a way to *represent* any multi-step from t by a sub-step of t^*.

Definition 2.21 Let ϕ be a multi-step from t. Its *canonisation* $\iota(\phi)$ is the multi-step obtained by replacing any redex-pattern ψ in ϕ by $\iota(\psi)$ (see Figure 7). A multi-step ϕ is called *canonical* if $\phi = \iota(\phi)$.

Example 2.22 We illustrate canonisation using the term t of Example 2.20. Canonising the step $\phi{:}\overline{P(S(}P(S(P(x))))) \mathrel{-\!\!\circ\!\!\to} P(S(P(x)))$ contracting the set $\{1\}$ yields the step $\iota(\phi){:}P(\underline{S(P(}S(P(x))))) \mathrel{-\!\!\circ\!\!\to} P(S(P(x)))$, contracting the set $\{2\}$ since the latter redex-pattern is contracted by t^* and its head

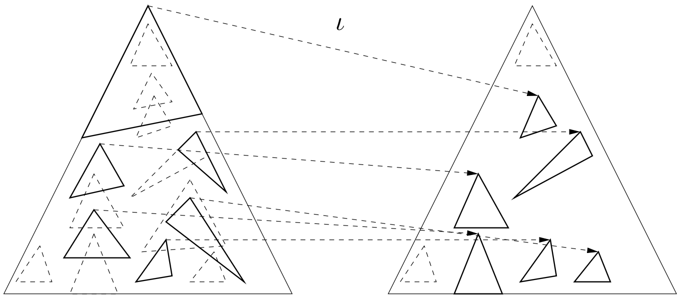

Figure 7. Multi-step ψ from t (of Figure 6) and its canonisation $\iota(\psi) \subseteq t^*$

has overlap with the former. Similarly, the canonisation of the multi-step $\psi{:}\overline{P(S}(\overline{P(S}(P(x))))) \mathrel{-\!\!\circ\!\!\rightarrow} P(x)$ contracting $\{1, 2\}$ is t^*.

Lemma 2.23 *Let $\phi{:}t \mathrel{-\!\!\circ\!\!\rightarrow} s$ be a multi-step in a weakly orthogonal PRS. Then*

- *canonisation is equipollent: $\#\iota(\phi) = \#\phi$,*
- *canonisation is invariant: $\iota(\phi){:}t \mathrel{-\!\!\circ\!\!\rightarrow} s$,*
- *canonisation embeds: $\iota(\phi) \subseteq t^*$, and*
- *canonisation is idempotent: $\iota(\iota(\phi)) = \iota(\phi)$.*

Proof

- As established above ι is a total function on all redexes in t. By the assumption that ϕ is a multi-step, distinct redex-patterns ψ, χ contracted by ϕ do not have overlap, so the head symbols of their canonisations $\iota(\psi)$, $\iota(\chi)$ must be distinct, hence their redex-patterns do not have overlap either and ι is total and injective on the redexes contracted by ϕ. Thus the multi-step ϕ and $\iota(\phi)$ are in bijective correspondence, i.e. they have the same cardinality.
- Invariance follows from the previous item by weak orthogonality, since any single step ψ has overlap with its canonisation $\iota(\psi)$.
- Embedding follows immediately per construction.
- By the previous item as canonisation is the identity on sub-steps of t^*.

□

The last item of the lemma expresses that canonisation does what it purports to do; it yields canonical multi-steps. Canonising $\phi{:}t \mathrel{-\!\!\circ\!\!\rightarrow} s$ coincides with so-called *orthogonalisation* of ϕ with respect to t^* [17, Prop. 8.8.23].

Lemma 2.24 *$\mathrel{-\!\!\circ\!\!\rightarrow}_{\mathcal{H}}$ has projection, if $\mathcal{H}$ is a weakly orthogonal PRS.*

Proof We define for arbitrary multi-steps $\phi{:}t \mathrel{-\!\!\circ\!\!\rightarrow} s$ and $\psi{:}t \mathrel{-\!\!\circ\!\!\rightarrow} u$, the residual of ϕ after ψ via the orthogonal projection / for PRSs and canonisation:

$$\phi /\!\!/ \psi = \iota(\iota(\phi)/\iota(\psi))$$

This is well-defined, since by Lemma 2.23, it holds $\iota(\phi){:}t \mathrel{-\!\!\circ\!\!\rightarrow} s$, $\iota(\psi){:}t \mathrel{-\!\!\circ\!\!\rightarrow} u$, and $\iota(\phi), \iota(\psi) \subseteq t^*$. Since / was established to be a projection, we conclude.□

Example 2.25 Consider the term $f(f(g(f(f(x)))))$ in the weakly orthogonal TRS with rules:

$$\begin{aligned} f(f(x)) &\to f(x) \\ g(x) &\to f(x) \end{aligned}$$

and the three steps from it: ϕ contracting the outermost ff-redex, ψ the g-redex, and χ the innermost ff-redex. Each of these steps is canonical, hence projecting them consists in projecting them as in the orthogonal case:

$$\begin{aligned} \phi/\psi &: \underline{f(f}(f(f(f(x))))) \mathrel{-\!\!\circ\!\!\rightarrow} f(f(f(f(x)))) \\ \chi/\psi &: f(f(f(\underline{f(f}(x))))) \mathrel{-\!\!\circ\!\!\rightarrow} f(f(f(f(x)))) \\ \phi/\chi &: \underline{f(f}(g(f(x)))) \mathrel{-\!\!\circ\!\!\rightarrow} f(g(f(x))) \\ \psi/\chi &: f(f(\underline{g}(f(x)))) \mathrel{-\!\!\circ\!\!\rightarrow} f(f(f(f(x)))) \end{aligned}$$

followed by canonisation of the resulting steps. As it turns out, each is already canonical except for the first one. Canonising the latter yields the single step $\iota(\phi \mathbin{/\!\!/} \psi) : f(\underline{f(f}(f(f(x))))) \mathrel{-\!\!\circ\!\!\rightarrow} f(f(f(f(x))))$. Using these, we may compute further projections, e.g.

$$\begin{aligned} &(\phi \mathbin{/\!\!/} \psi) \mathbin{/\!\!/} (\chi \mathbin{/\!\!/} \psi) : f(f(\underline{f(f}(x)))) \mathrel{-\!\!\circ\!\!\rightarrow} f(f(f(x))) \\ &(\phi \mathbin{/\!\!/} \chi) \mathbin{/\!\!/} (\psi \mathbin{/\!\!/} \chi) : \underline{f(f}(f(f(x)))) \mathrel{-\!\!\circ\!\!\rightarrow} f(f(f(x))) \end{aligned}$$

Remark 2.26 As witnessed by the previous example, the order in which residuals are taken matters: first projecting ϕ over ψ and then over (the residual of) χ may be different from first projecting ϕ over χ and then over (the residual of) ψ. Therefore, the so-called cube law as formulated in [17, Def. 8.7.2] for an arbitrary notion of projection $/$:

$$(\phi/\psi)/(\chi/\psi) = (\phi/\chi)/(\psi/\chi)$$

fails when taking $/\!\!/$ as the notion of projection for weakly orthogonal PRSs, whereas it does hold for $/$ in the orthogonal case. We do not know whether a residual system, extending the single step ARS, can be associated to a weakly orthogonal PRS or, for that matter, to a weakly orthogonal string rewriting system. If it could, it would simplify some of the constructions in Section 4.

3 Acyclicity for separate generalisations

In this section, we generalise the WN $\Rightarrow$ AC result, as obtained in [10] for orthogonal first-order term rewriting systems, into each of the three individual directions mentioned in the introduction:

(i) to the higher-order case, where terms may contain bound variables,

(ii) to the weakly orthogonal case, where rules may have (trivial) conflicts, and

(iii) to weak head normalisation (WHN), where terms have head normal forms.

We will do so by generalising the proof technique introduced in [10], the essence of which we will present now, after which we consider each of the generalisations in the subsequent sections. The essence of [10] is captured by the following two results and their proofs.

Theorem 3.1 (Head normalisation) *If a term allows a reduction containing infinitely many head steps in an orthogonal TRS, then it does not have a head normal form.*

Proof Using the terminology of [17, Defs. 9.1.12 and 9.1.18], the contrapositive of the statement of the theorem can be concisely expressed as: head strategies are hyper-head-normalising for orthogonal TRSs. To conclude, note that a head step contracts the unique outermost redex, which therefore is external [17, Lem. 9.2.34], and hence is head needed (see the text below [17, Def. 9.2.52]). Thus the result follows from hyper-head-normalisation of head needed strategies [17, Thm. 9.2.60]. □

Theorem 3.2 *Weakly normalising orthogonal TRSs are acyclic.*

Proof For a proof by contradiction, consider a term t in a weakly normalising orthogonal TRS $\mathcal{R}$ of minimal height which is cyclic. That is, there exists a non-empty reduction $\sigma{:}t \twoheadrightarrow t$. By the minimality assumption σ must contain some head step, hence the infinite reduction σ^ω from t obtained by repeating σ infinitely often contains infinitely many head steps. Therefore, by the Head Normalisation Theorem, t does not have a head normal form, hence *a fortiori* also no normal form, contradicting the assumption that $\mathcal{R}$ was weakly normalising. □

Note that for the proof of the latter theorem only the Head Normalisation Theorem and *closure under sub-terms* are needed. Hence, in the subsequent sections dealing with each of the individual extensions, to generalise Theorem 3.2, it suffices to generalise these two results.

3.1 Acyclicity of weakly normalising orthogonal higher-order *TRSs*

As higher-order term rewriting systems we take the PRSs of Subsection 2.3. To generalise the WN ⇒ AC result, we could use (see the proof of Theorem 3.1) hyper-head-normalisation of head needed strategies. Instead, in anticipation of the generalisations in Section 4, we will base our generalisation on head normalisation of so-called outermost-fair strategies [18, Thm. 2], where a strategy is *outermost-fair* [17, Def. 9.3.1] if each outermost redex is fairly treated in the sense that it will either be eventually contracted or its residual ceases to be outermost. Since this result only holds for PRSs which are *fully-extended*, meaning that whether a rule is applicable or not does not depend on whether a bound variable occurs in one of its arguments or not, we first establish the result for these.

Example 3.3 The η-rule of the λ-calculus is *not* fully-extended as its applicability depends on whether the bound variable occurs. In general, rule schemes containing so-called variable conditions yield non-fully-extended PRS rules.

Theorem 3.4 (Fully-extended higher-order head normalisation) *Head strategies are hyper-head-normalising for orthogonal fully-extended PRSs.*

Proof We have to prove head normalisation of any hyper-head-strategy, i.e. of any strategy which eventually always performs a head step. Now note that any hyper-head-strategy is obviously outermost-fair since at the moment a head redex is contracted it is the only outermost redex, and since the strategy is hyper-head, this will always eventually be the case. Thus the result follows from head normalisation of outermost-fair strategies for orthogonal fully-extended PRSs [18, Thm. 2]. □

Theorem 3.5 *A weakly normalising fully-extended orthogonal PRS is acyclic.*

Proof As for Theorem 3.2, using Theorem 3.4 and noting that reductions in fully-extended PRSs are closed under sub-terms. □

Example 3.6 The PRS representing the $\lambda\beta$-calculus (Example 2.13) is fully-extended and orthogonal. Hence we conclude that if it were normalising, then it would be acyclic. However, obviously, the implication is trivial since the $\lambda\beta$-calculus is *not* weakly normalising. Still, from the proof of Theorem 3.2 one notes that the result generalises to any *sub-calculus* of the λ-calculus which is closed under reduction and taking sub-terms.

For instance, typed $\lambda\beta$-calculi usually do have these properties (closure under reduction is there known as *subject reduction*), so for these one may conclude AC from their weak normalisation.

Recall from [4,10] that the *family* of a term in a term rewriting system, consists of all terms reachable from it w.r.t. the union of the reduction and the sub-term relations. Since 'sub-term steps' can be postponed, the same notion is obtained by considering only sub-terms of reachable terms. Hence, an arbitrary term induces a sub-calculus: its family.

That outermost-fair reductions need not be head normalising for non-fully extended orthogonal PRSs, is shown by the following example.

Example 3.7 Consider the term $t = f(x.g(e(x)))$ in the orthogonal PRS with rules

$$\begin{aligned} Z.f(x.Z) &\to Z.a \\ Z.g(Z) &\to Z.g(Z) \\ Z.e(Z) &\to Z.b \end{aligned}$$

with Z of base type, everywhere. The f-rule is non-fully-extended as it tests for the absence of the bound variable x in its argument Z. Since this test fails in t, the term itself is not yet a redex, and hence the g-redex is the outermost-redex, yielding an outermost-fair reduction cycle. However, a reduction to the head normal form a is possible, by first contracting the e-redex in t which *erases* the variable x, causing the test for the f-redex to succeed.

Note that the example is not a counter-example to the implication WN $\Rightarrow$ AC, since the sub-term $g(e(x))$ of t does not have a (head) normal form. Indeed, it could not be, as we will show now. The problem in the example is that the reduction step erasing the variable is 'needed' for the creation of the head redex-pattern, but does not 'contribute' to it via the ordinary descendant relation. We now construct the full-extension of a PRS $\mathcal{H}$ where this problem is resolved, while preserving the good properties of $\mathcal{H}$,

Definition 3.8 The *full-extension* $\bar{\mathcal{H}}$ of a PRS $\mathcal{H}$ is obtained by taking as rules the rules of $\mathcal{H}$, but supplying all 'missing' bound variables to the meta-variables in the left-hand side of each rule, and supplying arbitrary closed terms in the right-hand sides. Then a *redex-pattern* for rule $\bar{l} \to \bar{r}$ is defined to be an instance of $\bar{l}$ which can be reduced to an instance of l by means of $\mathcal{H}$-steps in its arguments. [7] All redex-pattern-defined notions for a PRS $\mathcal{H}$ are generalised to its full-extension $\bar{\mathcal{H}}$ via this notion of redex-pattern (formally, we consider consider the generated sub-PRS).

Clearly, every $\mathcal{H}$ redex-pattern, say for rule $l \to r$ is a $\bar{\mathcal{H}}$ redex-pattern for the corresponding fully-extended rule $\bar{l} \to \bar{r}$, and applying either then yields the same result since the closed terms supplied in the right-hand side $\bar{r}$ of the latter will be erased. In general, the $\mathcal{H}$-steps in arguments in the definition of $\bar{l}$ redex-pattern serve to erase variables which were restricted by variable conditions in $\mathcal{H}$.

Example 3.9 Fully-extending the PRS of Example 2.17 representing the $\lambda\beta\eta$-calculus, does not change the β-rule. However, the η-rule changes into

$$\bar{M}{:}\mathsf{Trm}{\to}\mathsf{Trm}.\mathsf{abs}\, x{:}\mathsf{Trm}.\mathsf{app}(\bar{M}x)x \to \bar{M}{:}\mathsf{Trm}{\to}\mathsf{Trm}.\bar{M}a$$

for an arbitrary closed term a. Switching back to ordinary λ-calculus notation again, the term $\lambda x.KIxx$ then is a redex-pattern for the $\bar{\eta}$-rule, since β-reducing its argument KIx to I yields the instance $\lambda x.Ix$ of the η-rule. However, $\lambda x.xIKx$ is not such a redex-pattern, since its argument xIK cannot be $\lambda\beta\eta$-reduced to a term not containing x (it is in normal form).

For an example from arithmetic, consider the term $t = \prod_{i=1}^{10} 5 + (0 \times i)$ in

$$0 \times x \to 0$$
$$\prod_{i=1}^{n} Z \to n \times Z$$

where the variable condition that i does not occur in Z, is brought about by

[7] This makes redex-patternhood undecidable. That does not matter here.

not supplying i as an argument to the meta-variable Z. We have

$$\prod_{i=1}^{10} 5 + (0 \times i) \to \prod_{i=1}^{10} 5 + 0 \to 10 \times (5 + 0)$$

as only reduction. Fully-extending the PRS leaves the first rule unchanged, but modifies the second rule into

$$\prod_{i=1}^{n} Z(i) \to n \times Z(0)$$

where we have (arbitrarily) supplied the constant 0 to the meta-variable Z in the right-hand side. In the full-extension t is a redex for this rules since $5 + (0 \times i)$ can be reduced to $5 + 0$ in which i does not occur, so we have another reduction (with the same result!)

$$\prod_{i=1}^{10} 5 + (0 \times i) \to 10 \times (5 + (0 \times 0)) \to 10 \times (5 + 0)$$

Lemma 3.10 *Let $\bar{\mathcal{H}}$ be the full-extension of the PRS $\mathcal{H}$. Then*

- $\to_{\mathcal{H}} \subseteq \to_{\bar{\mathcal{H}}}$,
- *the sets of (head) normal forms of $\mathcal{H}$ and $\bar{\mathcal{H}}$ coincide,*
- *$\bar{\mathcal{H}}$ is orthogonal if and only if $\mathcal{H}$ is.*
- *residuals of $\bar{\mathcal{H}}$-redex-patterns are $\bar{\mathcal{H}}$-redex-patterns again.*
- *$\bar{\mathcal{H}}$-reductions are closed under standardisation.*

Proof

- It suffices to note that every $\mathcal{H}$-redex-pattern is a $\bar{\mathcal{H}}$-redex-pattern.
- Although a term t may have a (head) redex-pattern for $\bar{\mathcal{H}}$ but not for $\mathcal{H}$, that may exist only if t may be further reduced to a (head) redex-pattern in $\mathcal{H}$ proving the result, using the previous item.
- Since every $\mathcal{H}$-redex-pattern is a $\bar{\mathcal{H}}$-redex-pattern, the 'if-direction' is trivial. To prove the 'only-if-direction', consider for a proof by contradiction among the set of redex-patterns having overlap two overlapping redex-patterns ϕ, ψ such that the (sum of their) lengths of the witnessing $\mathcal{H}$-multi-step-reductions is minimal, and among those such that the (sum of the) number of contracted redex-patterns in their first steps is minimal. Consider a redex-pattern χ of maximal depth in the first multi-steps of these reductions, say w.l.o.g. it is in the witnessing reduction for ϕ, i.e. it is in some argument of ϕ. We distinguish cases on the relative position of χ w.r.t. ψ.

  If χ is parallel to or entirely below ψ, then we obtain a contradiction with minimality by contracting χ, and considering the residuals of ϕ and ψ after χ, together with the residuals of the witnessing reductions after χ. The latter have at most the same length, but if they have the same length then the number of contracted redex-patterns in their first steps has become smaller by choice of χ.

If χ overlaps ψ, then we obtain a contradiction by considering the redex-patterns ψ and χ and their respective witnessing reductions, where the length of the latter is 0, by χ being a $\mathcal{H}$-redex-pattern, so certainly less than the witnessing reduction of ϕ.

If χ is entirely above ψ, then we obtain a contradiction with the assumptions that ϕ and ψ have overlap and that χ is entirely below ϕ.

Since these are all possibilities and all of them yield a contradiction, we conclude that both witnessing reductions must be empty, hence ϕ and ψ are in fact overlapping $\mathcal{H}$-redex-patterns, contradicting orthogonality of $\mathcal{H}$.

- Any residual of a $\bar{\mathcal{H}}$-redex-pattern ϕ is a substitution instance of ϕ. Hence, by performing the same reduction on the residual of ϕ, the same variables bound in the pattern will be erased.
- Let $\phi \cdot \psi$ be an anti-standard pair, starting from t in a non-standard reduction, i.e. a pair such that either ϕ and ψ are parallel or ψ is entirely outside ϕ. Standardisation consists in permuting such anti-standard pairs. It suffices to show that permuting the redex-patterns ϕ and ψ yields redex-patterns again. By construction of the full-extension of a PRS, the origin ψ' of the redex-pattern ψ in t is a redex-pattern again, since by definition it is outside or parallel to ϕ. In the latter case the property follows since the redex ψ and its origin ψ' are identical. In the former case, it suffices to prefix ϕ, which is then a $\mathcal{H}$-step in an argument of ψ', to the reduction which witnessed that ψ was a redex-pattern. We conclude since the residual(s) of ϕ after ψ' are redex-patterns again by the previous item.

□

Theorem 3.11 (Higher-order head normalisation) *Head strategies are hyper-head-normalising for orthogonal PRSs.*

Proof Let $\bar{\mathcal{H}}$ be the full-extension of the orthogonal PRS $\mathcal{H}$. By Lemma 3.10, a hyper-head-reduction for $\mathcal{H}$ is a hyper-head-reduction for $\bar{\mathcal{H}}$ as well. By Theorem 3.4, therefore the $\mathcal{H}$-reduction will find a $\bar{\mathcal{H}}$-head-normal form if one exists. Therefore, by Lemma 3.10 again, the reduction will find a $\mathcal{H}$-head-normal form as well.

The only further point to consider is that Theorem 3.4 holds for fully-extended orthogonal PRSs. Thereto, we have to check that its proof goes through for the sub-PRS $\bar{\mathcal{H}}$. Since the proof was based on [18, Thm. 2] stating that outermost-fair reductions are head normalising for fully-extended orthogonal PRSs, it suffices to show that theorem goes through for $\bar{\mathcal{H}}$. Apart from orthogonality, its proof requires the following properties, all of which are implied by Lemma 3.10: [8]

- residuals of redex-patterns are redex-patterns again,

[8] The last two properties fail for $\mathcal{H}$.

- head normal forms are closed under reduction,
- reductions are closed under standardisation, and
- head normal forms are closed under expansions below the head.

□

Using this theorem, we obtain the desired generalisation of Theorem 3.5

Theorem 3.12 *Weakly normalising orthogonal PRSs are acyclic.*

Example 3.13 The $\lambda\eta$-calculus (having *only* the η-rule, *not* the β-rule), is an orthogonal PRS, hence we may conclude its acyclicity from its weak normalisation (it is even strongly normalising, but that's not the point here).

3.2 Acyclicity of weakly normalising weakly *orthogonal TRSs*

The generalisation of WN $\Rightarrow$ AC from orthogonal to weakly orthogonal TRSs is entirely unproblematic.

Theorem 3.14 (Weakly orthogonal head normalisation) *Head strategies are hyper-normalising for weakly orthogonal TRSs.*

Proof As for Theorem 3.4, but using normalisation of outermost-fair strategies for *weakly* orthogonal TRSs [18, Thm. 1], instead of *head* normalisation of outermost-fair strategies for orthogonal TRSs [18, Thm. 2]. □

Theorem 3.15 *Weakly normalising weakly orthogonal PRSs are acyclic.*

Example 3.16 The predecessor/successor TRS of Example 2.17 is weakly orthogonal, so its acyclicity may be concluded from its weak normalisation.

3.3 Acyclicity of weakly head *normalising orthogonal TRSs*

The generalisation of the implication WN $\Rightarrow$ AC to WHN $\Rightarrow$ AC is trivial, noting that the proof of the former already established the latter.

Theorem 3.17 *Weakly head normalising orthogonal TRSs are acyclic.*

Proof As for Theorem 3.2, skipping the proof step that non-head-normalisation implies non-normalisation. □

4 Acyclicity for the pair-wise combinations

We now consider the pair-wise combinations of the generalisations of the previous sections. The proofs of two of them are straightforward adaptations of the proofs there. The third combination turns out to be much more subtle and will take up the rest of this section, the longest of the paper.

Theorem 4.1 *For weakly orthogonal fully-extended PRSs, WN $\Rightarrow$ AC.*

Proof As for Theorem 3.15, noting that the normalisation result on which it was based ([18, Thm. 1]), holds not only for TRSs, but for fully-extended PRSs. □

Remark 4.2 We don't know whether the fully-extendedness condition can be dropped in this case, but see Theorem 5.3 for the special case of the (non-fully-extended) $\lambda\beta\eta$-calculus. Two problems with the proof method of the previous section are that the full-extension of a weakly orthogonal PRS (see Definition 3.8) need not be weakly orthogonal, and that head normal forms in the full-extension need not be closed under expansions below the head.

Theorem 4.3 *For orthogonal PRSs, WHN $\Rightarrow$ AC.*

Proof As for Theorem 3.12, but again simpler since the implication from non-head-normalisation to non-normalisation need not be used. □

Example 4.4 Weakly head-normalising sub-calculi of the $\lambda\beta$-calculus are acyclic. See Example 3.6.

We now turn our attention to the third combination, weak head normalisation with weak orthogonality. In particular, we show that reductions are acyclic in weakly orthogonal TRSs which are weakly head normalising. The proof will be based on projecting cyclic reductions onto cyclic reductions. More precisely, we first show that if a term t would allow both a cyclic reduction σ and a multi-step reduction to head normal form s, then we may create (Definition 2.8) preservation of cyclicity for the steps of the latter. That is, we may assume that the residual of σ after each of the the respective multi-steps of the latter reduction, is cyclic. After that, we show that we may even assume that the projections preserve head-cyclicity, i.e. preservation of cycles which contains a(t least one) head step. This then yields a contradiction, since σ would project onto a head-cycle from s, which was assumed to be in head normal form.

4.1 Preservation of cycles

We first observe that the projection of a cyclic reduction over a (multi-)step does not yield, in general, a cyclic reduction again. To overcome this problem, we then show that we may create this property in the sense of Definition 2.8. That is, we can always extend the multi-step to another multi-step which does preserve cyclicity. By Lemma 2.9, this suffices to create this property for any reduction from t to head normal form, with respect to a given cycle on t.

That residuals of cycles after *arbitrary* multi-steps need *not* be cycles is witnessed by the following example.

Example 4.5 The step $\sigma{:}\underline{f}(a, a, a) \to f(a, a, a)$ is a cycle in the TRS

$$\begin{aligned} a &\to b \\ f(x, y, z) &\to f(y, x, a) \end{aligned}$$

However, the residual of σ after the parallel step $\phi{:}f(\underline{a}, a, \underline{a}) \mathrel{-\!\!\!+\!\!\!\!\rightarrow} f(b, a, b)$ is $\underline{f}(b, a, b) \to f(a, b, a)$ which obviously is *not* a cycle (see Figure 8 left).

$$\begin{tikzcd} f(a,a,a) \arrow[r, "\sigma"] \arrow[d, "\phi"'] & f(a,a,a) \arrow[d] \\ f(b,a,b) \arrow[r, "\text{acyclic}"'] & f(a,b,a) \end{tikzcd} \qquad \begin{tikzcd} f(a,a,a) \arrow[r, "\sigma"] \arrow[d, "\bar{\phi}"'] & f(a,a,a) \arrow[d, "\bar{\phi}"] \\ f(b,b,a) \arrow[r, "\text{cyclic}"'] & f(b,b,a) \end{tikzcd}$$

Figure 8. Residuals of the cycle σ along the step ϕ and its extension $\bar{\phi}$, respectively.

Remark 4.6 The phenomenon of non-preservation of reduction cycles is well-known from infinitary rewriting [17, Ex. 12.3.3]. There, it is used to illustrate the fact that the notion of Cauchy convergence does not agree well with projection, i.e. residual theory.

However, there always is an extension of the step which *does* preserve cyclicity.

Example 4.7 The parallel step $\bar{\phi}{:}f(\underline{a}, \underline{a}, a) \mathrel{-\!\!\!+\!\!\!\!\rightarrow} f(b, b, a)$ is an extension of the parallel step ϕ of Example 4.5, since the target $f(b, a, b)$ of ϕ reduces to the target $f(b, b, a)$ of $\bar{\phi}$ as one easily checks. Note that the residual of the cycle σ after $\bar{\phi}$ is $\underline{f}(b, b, a) \to f(b, b, a)$ which *is* a cycle (see Figure 8 right).

The problem is to find an *algorithm* for constructing such a $\bar{\phi}$ from ϕ. That is, which constructs for an arbitrary step ϕ co-initial to some cycle σ, an extension $\bar{\phi}$ of ϕ which preserves cyclicity, i.e. such that the residual $\sigma/\bar{\phi}$ of σ after $\bar{\phi}$ is a cycle again. We will present, in two stages, an algorithm for doing just that for multi-steps.

The orthogonal case is treated first since it not only provided us with the insight used in our particular definition of the weakly orthogonal projection in Section 2.4 above, but it also brings to the fore a, potentially interesting, correspondence between projecting cycles in residual theory and dividing natural numbers in arithmetic.

4.1.1 Preservation of cycles in orthogonal PRSs

First, we show that for a given cycle σ and co-initial multi-step ϕ, there is an extension ϕ_0 of ϕ such that some *repetition* of σ projects onto a cycle, when projected over ϕ_0. Next, we show that all the residuals of ϕ_0 along this repetition can be *compressed* into a multi-step $\bar{\phi}$, which extends ϕ_0, such that σ *itself* projects onto a cycle when projected over $\bar{\phi}$.

Consider repeatedly projecting the step ϕ in Example 4.5 over the cycle σ (see Figure 9). First, projecting ϕ over σ yields $\phi_0{:}f(a, \underline{a}, a) \mathrel{-\!\!\!+\!\!\!\!\rightarrow} f(a, b, a)$. Next, projecting ϕ_0 over σ yields $\phi_1{:}f(\underline{a}, a, a) \mathrel{-\!\!\!+\!\!\!\!\rightarrow} f(b, a, a)$. Finally, projecting ϕ_1 over σ yields ϕ_0 again. That is, the series of residuals eventually becomes repetitive. This situation is completely general.

$$\begin{tikzcd}
f(a,a,a) \arrow[r, ->>, "\sigma"] \arrow[d, "\phi"'] & f(a,a,a) \arrow[r, ->>, "\sigma"] \arrow[d, "\phi_0"] & f(a,a,a) \arrow[r, ->>, "\sigma"] \arrow[d, "\phi_1"] & f(a,a,a) \arrow[r, ->>, "\sigma"] \arrow[d, "\phi_0"] & {} \\
f(b,a,b) \arrow[r, ->>] & f(a,b,a) \arrow[r, ->>] & f(b,a,a) \arrow[r, ->>] & f(a,b,a) \arrow[r, ->>] & {}
\end{tikzcd}$$

Figure 9. Repetition in computing residuals of ϕ after the cycle σ.

Lemma 4.8 (Repetition) *Let ϕ be a multi-step co-initial to a reduction cycle σ in an orthogonal PRS. Then there exists a multi-step ϕ_0 extending ϕ, and a positive natural number n, such that $\phi_0/\sigma^n = \phi_0$.*

Proof By Lemma 2.15, the multi-step ARS $\multimap$ associated to an orthogonal PRS has projection, say it is denoted by $/$. Consider the infinite sequence of residuals ϕ/σ^i of ϕ after the i-fold repetition of i, for arbitrary i. Since redex-patterns in a term in an orthogonal TRS do not have overlap, $\multimap$ is finitely branching. Therefore, by the Pigeon Hole Principle, the sequence of residuals must be eventually repetitive. Moreover, the target of ϕ reduces to the target of any residual in the sequence, since the target of ϕ/σ^i reduces to the target of ϕ/σ^{i+1} via reduction $\sigma/(\phi/\sigma^i)$. Hence, taking the first residual on the repetition yields the desired extension. $\square$

The extension of ϕ constructed in the proof of the lemma will be called the *repetitive* extension of ϕ and denoted by ϕ_0. It is easy to see that in general it may take an arbitrary number of cycles before repetition sets in, and that the repetition itself may also take an arbitrary number of cycles. For instance, taking the residual of the parallel step $f(a,a,a,a,a,a) \mathrel{+\!\!\!\!\rightarrow} f(b,a,a,b,a,a)$ for the rule $a \to b$, after the single-step cycle with respect to the rule $f(x_1,x_2,x_3,z_1,z_2,z_3) \to f(x_2,x_3,x_1,a,z_1,z_2)$ takes 3 cycles to become repetitive with a repetition of length 3.

Remark 4.9 If, in the lemma, the cycle σ eliminates (either contracts or erases) all the redexes of ϕ, then ϕ_0 is the empty multi-step. This is witnessed e.g. by the parallel step $\phi{:}f(\underline{a},\underline{a}) \mathrel{+\!\!\!\!\rightarrow} f(b,b)$ and the cycle $\sigma{:}f(\underline{a},a) \to \underline{f}(b,a) \to f(\underline{b},b) \to f(a,\underline{b}) \to f(a,a)$, for the TRS with rules $a \to b$, $b \to a$ and $\underline{f}(x,y) \to f(x,x)$.

Dually, if the multi-step ϕ erases all steps on the cycle σ, then the residual of the cycle is empty. This is witnessed e.g. by the step $\phi{:}\underline{f}(a,b) \to f(a,a)$ and the cycle $\sigma{:}f(a,\underline{b}) \to f(a,\underline{a}) \to f(a,b)$ for the same rules.

Remark 4.10 The situation is similar to what happens in arithmetic when dividing two natural numbers: the resulting decimal expansion eventually becomes repetitive, after an initial part of arbitrary length. Figure 10 illustrates this for dividing 45 by 33, which results, after an initial digit 1, into a repetition of 36. The reason for the decimal expansion being repetitive is indeed the same as for residuals: since there are only finitely many distinct numbers modulo 33, the Pigeon Hole Principle yields that eventually a remainder,

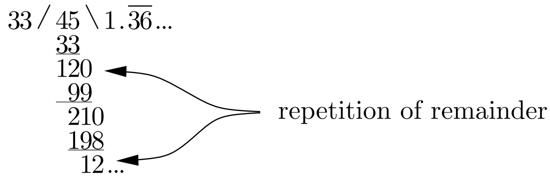

Figure 10. Repetition in computing decimal expansion when dividing by 33

i.e. residual, reoccurs (in the example 12 reoccurs). Since the algorithm is deterministic, the decimal expansion is repetitive from there on.

The extension process is monotonic.

Proposition 4.11 *Let σ be a cycle co-initial to both ϕ and ψ and let $\phi \subseteq \psi$, i.e. the set of redexes contracted by ϕ is contained in the set of redexes contracted by ψ. Then ψ_0 is an extension of ϕ_0.*

Proof By the Repetition Lemma the situation must be as displayed in Figure 11: After a number of σ-cycles both the residuals of ϕ and ψ become

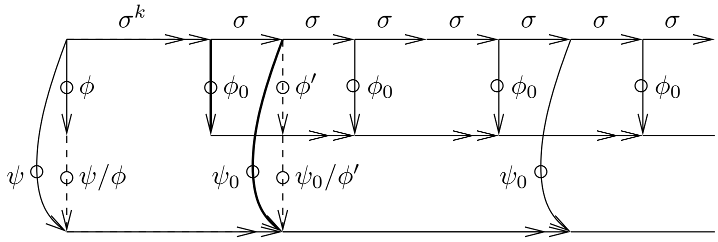

Figure 11. Monotonicity of extensions

repetitive, for extensions ϕ_0 and ψ_0, respectively. However, the number of σ-cycles required for either repetition may differ (in the Figure they are 2 and 4, respectively) and moreover, the repetitions may or may not set in at the same moment (in the Figure they do not). Now consider the first ψ_0 which is on a ϕ_0-repetition. In the figure this is the thick curved arrow, and we can immediately see that it is an extension of ϕ_0 since its target is reachable from the target of the thickly drawn occurrence of ϕ_0 to its left.

Formally, we reason as follows. By the Repetition Lemma: $\phi_0 = \phi/\sigma^{n'}$, $\phi_0/\sigma^n = \phi_0$, $\psi_0 = \psi/\sigma^{m'}$, and $\psi_0/\sigma^m = \psi_0$. for some natural numbers n', m' and positive natural numbers n, m. By Archimedeanity, there is some k such that $n' \leq m' + km$. We check that $\psi/\sigma^{m'+km}$ proves the lemma using algebra for residual systems [17]:

- 'it is the thick curved arrow': $\psi/\sigma^{m'+km} = (\psi/\sigma^{m'})/\sigma^{km} = \psi_0/\sigma^{km} = \psi_0$.
- 'the thickly drawn occurrence of ϕ_0 is to its left': the target of ϕ_0 reduces to the target of $\phi/\sigma^{m'+km}$ (ϕ' in Figure 11) as $\phi_0 = \phi/\sigma^{n'}$ and $n' \leq m' + km$.
- 'the target is reachable from the thickly drawn occurrence of ϕ_0 to its left': by the previous item, it remains to show that the target of $\psi/\sigma^{m'+km}$ is

reachable from the target of $\phi/\sigma^{m'+km}$. We claim the more general fact that the target of ψ/σ^i is reachable from the target of ϕ/σ^i, for any i.

The claim holds for any residual system with union. Firstly, union distributes over residuation in such systems, so ϕ/σ^i is a sub-step of ψ/σ^i:

$$\psi/\sigma^i = (\phi \cup \psi)/\sigma^i = (\phi/\sigma^i) \cup (\psi/\sigma^i)$$

Secondly, for any sub-step χ of a multi-step ζ, it holds by general residual theory (e.g. using the finite developments theorem), that ζ has the same source and target as the reduction consisting of χ followed by ζ/χ,[9] so ζ extends χ. Applying this to the sub-step ϕ/σ^i of ψ/σ^i proves the claim.

Since extension is a quasi-order as observed above, combining the three items yields that ψ_0 is an extension of ϕ_0. □

Remark 4.12 For ϕ and ψ as in the proof, it holds that the length of either their repetitions may exceed that of the other. For instance, let ϕ contract the first a in $f(a,a,a,a,a)$, and let ψ be the parallel step contracting the third a as well, in a TRS with rules $a \to b$ and $f(x_1,x_2,y_1,y_2,y_3) \to f(x_2,x_1,y_2,y_3,y_1)$. For the single step cycle σ with respect to the latter rule, we have $\phi_0 = \phi$ and $\psi_0 = \psi$, with respective repetition lengths 2 and 6. However, if we take for ψ the parallel step contracting the first two as, its repetition has length 1.

Note that in the former example, the 'generators' are on distinct cycles, hence the length of their combined repetition is a multiple (in fact, the least common multiple) of their individual repetition lengths. The latter example combines two distinct generators of the same cycle yielding the opposite situation that the length of their combined repetition divides the lengths (which are both 2) of their individual repetition.

As an immediate consequence of the Repetition Lemma, projecting *some repetition* σ^n of σs over the extension ϕ_0 of some multi-step ϕ, is a cycle again. By monotonicity we now also know that the extension t_0^* of the maximal multi-step t^* from t, extends all such ϕ_0. However, the goal was to find an extension $\bar{\phi}$ of ϕ such that projecting σ *itself* over $\bar{\phi}$ yields a cycle. This we do by *compressing* all the residuals of the multi-step ϕ_0 occurring along the repetition of σ, into a single multi-step.

To see how compression works, reconsider Figure 9, showing a repetition of length two. For the residuals ϕ_0 and ϕ_1 on the repetition it holds:

$$\phi_0/\sigma = \phi_1$$
$$\phi_1/\sigma = \phi_0$$

Combining these two equalities yields, by distributivity:

$$(\phi_0 \cup \phi_1)/\sigma = (\phi_0/\sigma) \cup (\phi_1/\sigma) = \phi_1 \cup \phi_0 = \phi_0 \cup \phi_1$$

In other words, the union of the multi-steps on the repetition is its *own* residual after the cycle σ. Since the union $\bar{\phi} = \bigcup_{0 \le i \le 1} \phi_i$ contains ϕ_0, it extends ϕ_0,

[9] In fact, the reduction ζ is even *permutation equivalent* to $\chi \cdot (\zeta/\chi)$.

which in turn extends ϕ by the Repetition Lemma, from which we conclude preservation of cyclicity. This situation is completely general.

Lemma 4.13 (Compression) *Let $\phi/\sigma^n = \phi$, for some step ϕ co-initial to a cycle σ and positive natural number n, and let $\bar{\phi}$ be its (n-fold) compression defined as $\bigcup_{0\leq i<n} \phi_0/\sigma^i$. Then $\bar{\phi}/\sigma = \bar{\phi}$ and $\bar{\phi}$ is an extension of ϕ.*

Proof Using distributivity, we compute

$$\bar{\phi}/\sigma = (\bigcup_{0\leq i<n} \phi_i)/\sigma = \bigcup_{0\leq i<n} (\phi_i/\sigma) = \bigcup_{0\leq i<n} \phi_{i+1 \operatorname{\mathsf{mod}} n} = \bigcup_{0\leq i<n} \phi_i = \bar{\phi}$$

That $\bar{\phi}$ is an extension of ϕ follows from $\bar{\phi} = \bigcup_{0\leq i<n} \phi/\sigma^i \supseteq \phi/\sigma^0 = \phi$, as in the proof of Proposition 4.11. ☐

Remark 4.14 Although residuals of parallel steps are parallel steps again in first-order TRSs, their compression may yield a multi-step which is *not* parallel, as witnessed by the following example. Consider the term $t = f(c(c(c(a))), c(c(c(a))))$ in the orthogonal TRS with rules

$$\begin{aligned} a &\to c(a) \\ c(x) &\to x \\ f(x,y) &\to f(y, c(x)) \end{aligned}$$

The term t allows the cyclic reduction σ:

$$\begin{aligned} &f(c^1(c^2(\underline{c^3}(a))), c^4(c^5(c^6(a)))) \\ &\to f(c^1(c^2(a)), \underline{c^4}(c^5(c^6(a)))) \\ &\to f(c^1(c^2(a)), c^5(c^6(\underline{a}))) \\ &\to \underline{f}(c^1(c^2(a)), c^5(c^6(c(a)))) \\ &\to f(c^5(c^6(c(a))), c(c^1(c^2(a)))) \end{aligned}$$

where we have numbered the occurrences of c in t to ease tracing each along σ. It is easy to see that the parallel step ϕ_0 contracting the set $\{c^2, c^5\}$ of c-redexes projects onto the parallel step ϕ_1 contracting the set $\{c^6, c^1\}$, over one σ-cycle, which projects onto the original ϕ_0 after another σ-cycle. However, their compression

$$\bar{\phi} = \phi_0 \cup \phi_1$$

is the multi-step contracting $\{c^1, c^2, c^5, c^6\}$, hence is not parallel as c^1 nests c^2 (and c^5 nests c^6).

Remark 4.15 Continuing Remark 4.10, we note that also compression can be applied to repeating decimals. To see this, consider computing the decimal expansion of $\frac{6}{7}$. Obviously, for each pair r, r' of consecutive remainders it holds:

$$r \cdot 10 = r' \operatorname{\mathsf{mod}} 7$$

Now consider the 'compression' $\bar{r} = \Pi_{0\leq i<n} r_i$ of all the remainders along the repetition. Since multiplication *distributes* over taking the remainder modulo

7, combining the equations for all consecutive pairs of residuals yields:

$$\bar{r} \cdot 10^n = \bar{r} \,\mathsf{mod}\, 7$$

From this we may conclude, *without actually computing* $\bar{r}$, that either $\bar{r} = 0$ or $10^n = 1 \,\mathsf{mod}\, 7$, hence that the repetition is either $\overline{0}$ or has length 6. In case of $\frac{6}{7}$ the latter possibility holds: It computes to $0.\overline{857142}$ with remainders $r_0 = 4$, $r_1 = 5$, $r_2 = 1$, $r_3 = 3$, $r_4 = 2$, and $r_5 = 6$. Hence $n = 6$. That one can say something about the length of the repetition in a repeating decimal is well-known. The point here is the analogy between this result and the Compression Lemma. Of course, the analogy only goes so far. For instance, the compression $\bar{r} \,\mathsf{mod}\, 7$ ($= 720 \,\mathsf{mod}\, 7 = 6$) of all the remainders in $\frac{6}{7}$ is itself not repetitive; a change of base, from 10 to 10^6, is required for that.

Theorem 4.16 *Let ϕ be a multi-step co-initial to a cycle σ. Then there exists a multi-step $\bar{\phi}$ extending ϕ such that $\bar{\phi}/\sigma = \bar{\phi}$, hence $\sigma/\bar{\phi}$ is a cycle again.*

Proof By the Repetition Lemma, there exist an extension ϕ_0 of ϕ and a positive natural number n, such that $\phi_0/\sigma^n = \phi_0$. Defining $\bar{\phi}$ of ϕ as the n-fold compression of ϕ as in the Compression Lemma, we obtain by it that $\bar{\phi}$ is an extension of ϕ_0 for which it holds $\bar{\phi}/\sigma = \bar{\phi}$. The result follows by transitivity of extension. □

The extension $\bar{\phi}$ of ϕ constructed in the proof of the lemma will be called the *cyclic* extension of ϕ. For instance, the cyclic extension of the step ϕ of Example 4.5 is the one displayed on the right in Figure 8.

As an immediate consequence, the cyclic extension $\bar{\phi}$ of a multi-step ϕ witnesses that the multi-step ARS $\multimap\!\!\!\to$ creates preservation of cyclicity for an arbitrary cycle σ co-initial to ϕ (whence our naming of this extension).

Corollary 4.17 *Let t allow both a reduction cycle σ and a multi-step reduction τ to head normal form s in an orthogonal PRS. Then there is a multi-step reduction $\bar{\tau}$ from t to head normal form of the same length as τ, such that $\bar{\tau}/\sigma = \bar{\tau}$, so $\sigma/\bar{\tau}$ is cyclic.*

Proof By Theorem 4.16 and Lemma 2.9. □

We conclude this section with observing that for the maximal multi-step t^* from a term t, its repetitive and cyclic extensions coincide. That is, repeatedly taking residuals of t^* along σ eventually ends in a repetition of length 1. For instance, starting out with the maximal multistep $\underline{f}(\underline{a}, \underline{a}, \underline{a}) \multimap\!\!\!\to f(b, b, a)$ in Example 4.5, taking its residual along the cycle σ yields the multi-step $f(\underline{a}, \underline{a}, a) \multimap\!\!\!\to f(b, b, a)$, which has the desired properties. Before showing that this situation is completely general, we first show that also the cyclic extension is monotonic.

Proposition 4.18 *If $\phi \subseteq \psi$, then $\bar{\psi}$ extends $\bar{\phi}$.*

Proof As shown in the proof of Proposition 4.11, every residual of ϕ_0 is contained in some residual of of ψ_0. From this the result follows by monotonicity of union. We leave the formalisation to the reader. □

Thus, by monotonicity we know that the cyclic extension $\bar{t}^*$ of the maximal multi-step t^* from t extends the cyclic extension $\bar{\phi}$ of any multi-step from t. Using monotonicity, we show that its repetitive and cyclic extensions *coincide*.

Proposition 4.19 $t_0^* = \bar{t}^*$.

Proof We claim that the series of residuals of t^* along a co-initial cycle σ is monotonically decreasing. That is, $\phi/\sigma \subseteq \phi$, for any such residual ϕ. From this the proposition follows since t^* only consists of finitely many redexes, hence eventually the decreasing sequence must stabilise (quite possibly in the empty multi-step from t).

The claim is proven by induction on the index of ϕ in the series.

- If the index of ϕ is 0, then the claim trivially holds since $\psi \subseteq t^*$ for *any* multi-step ψ from t.
- Suppose the index of ϕ is $n+1$ and let ψ be the step with index n. Then $\psi/\sigma = \phi$ and, by the induction hypothesis, $\phi \subseteq \psi$, i.e. $\phi \cup \chi = \psi$ for some χ. Hence, by distributivity,
$$\phi/\sigma \subseteq (\phi/\sigma) \cup (\chi/\sigma) = (\phi \cup \chi)/\sigma = \psi/\sigma = \phi$$

□

4.1.2 Preservation of cycles in weakly orthogonal PRSs

We extend the results for orthogonal PRSs of the previous section to weakly orthogonal PRSs.

Lemma 4.20 *Let $\multimap$ be the multi-step ARS with projection $/\!\!/$. Then for its sub-ARS of canonical multi-steps it holds:*

- *the sub-ARS is finitely branching, i.e. an object is source of only finitely many canonical multi-steps, and*
- *projection $/\!\!/$ distributes over union.*

Proof

- By Lemma 2.23, any canonical multi-step from a term t is a sub-multi-step of the maximal multi-step t^* from t. The result follows since t^* being finite, it has only finitely many distinct sub-multi-steps.
- Let ϕ, ψ, χ be co-initial canonical multi-steps. Noting that canonisation distributes over union since it is defined redex-pattern-wise, we compute:
$$\begin{aligned}
&(\phi \cup \psi)/\!\!/\chi \\
&= \iota((\phi \cup \psi)/\chi) \\
&= \iota((\phi/\chi) \cup (\psi/\chi)) \\
&= \iota(\phi/\chi) \cup \iota(\psi/\chi) \\
&= (\phi/\!\!/\chi) \cup (\psi/\!\!/\chi)
\end{aligned}$$

□

Theorem 4.21 *Let ϕ be a multi-step co-initial to a cycle σ. Then there exists a multi-step $\bar{\phi}$ extending ϕ such that $\bar{\phi}/\!\!/\sigma = \bar{\phi}$, hence $\sigma/\!\!/\bar{\phi}$ is a cycle again.*

Proof It suffices to verify that the Repetition and Compression Lemmas still hold in the weakly orthogonal setting. First of all note that since the weakly orthogonal projection yields canonical multi-steps, we may restrict our attention in our generalisation to these.

Apart from requiring projection, the proof of the Repetition Lemma only requires that the canonical multi-step ARS be finitely branching. This was established in the first item of Lemma 4.20.

The proof of the Compression Lemma requires firstly that the union of a set of co-initial canonical multi-steps is again such a step. This holds, since sub-steps of the maximal multi-step are closed under union. Secondly, the proof requires distribution of projection over union, which was established in the second item of Lemma 4.20. □

Remark 4.22 Both the Repetition Lemma and the Compression Lemma can be entirely cast into the theory of abstract residual systems [17, Sec. 8.7]. For the former to hold, the abstract residual system needs to be finitely branching. For instance, it holds for parallel reduction in orthogonal TRSs. For the latter to hold, there must be a union (least upper bound) operation which distributes over residuation. Distribution is an algebraic law well-known both from the $\lambda\beta$-calculus [4] and from orthogonal TRSs [9], and holds in general for multi-steps in orthogonal rewriting systems. For instance, both the Repetition Lemma and the Compression Lemma hold for orthogonal graph rewriting systems such as interaction nets.

Corollary 4.23 *Let t allow both a reduction cycle σ and a multi-step reduction τ to head normal form s in a weakly orthogonal PRS. Then there is a multi-step reduction $\bar{\tau}$ from t to head normal form of the same length as τ, with $\bar{\tau}/\!\!/\sigma = \bar{\tau}$, so $\sigma/\!\!/\bar{\tau}$ is cyclic.*

Proof By Theorem 4.21 and Lemma 2.9. □

4.2 Preservation of parallel head cycles

Extending the results of the previous section, we show here that we may create the property that *parallel head* cycles are preserved. More precisely, we show that we may assume that the residual of a given cycle consisting of parallel steps of which at least one is a head step, is such a cycle again, or otherwise a non-empty parallel cycle which is shorter.

As before, we first present the orthogonal case in order to exhibit the overall proof structure, after which that is adapted to the, much more involved, weakly orthogonal case. Unlike before, we restrict ourselves to the first-order TRS case (parallel steps do not behave well in the higher-order case).

4.2.1 Preservation of parallel head cycles in orthogonal TRSs

We show that the property that 'parallel head cycles are preserved or give rise to non-empty parallel cycles which are shorter', is created by orthogonal TRSs. The proof will be based on the following two lemmas, the first of which (Head) deals with the case that the head step is preserved hence with preservation of parallel head cycles, and the second (Non-Head) deals with the case that the head step is eliminated.

The proof of the Head Lemma is trivial here, but in the weakly orthogonal case we will have to work hard to create the property, see Example 4.30.

Lemma 4.24 (Head) *Let ψ be a head step co-initial to a multi-step ϕ in an orthogonal TRS. Then its residual ψ/ϕ is either empty or a head step again.*

Definition 4.25 The measure $\mu(\phi)$ of a multi-step ϕ is the maximal length of an outermost development of ϕ, where the *outermost development* of a multi-step ϕ is defined by:

- If ϕ is the empty multi-step, the result is the empty reduction.
- Otherwise, the result is the contraction of an outermost redex ϕ_{om} followed by an outermost development of the multi-step ϕ/ϕ_{om}.

By the Finite Developments Theorem [17, Thm. 4.5.4] μ is well-defined. In fact, all outermost developments of a multi-step have the same length, which follows e.g. by observing that they are self-delimiting [19]. Outermost developments could also be called parallel standard developments, as they *are* those developments which are parallel standard [17, Def. 8.5.6].

Example 4.26 Consider the maximal multi-step $t^*{:}\underline{f}(\underline{f}(\underline{a},\underline{a}),\underline{f}(\underline{a},\underline{a})) \multimap g(g(b,b),g(b,b))$ from the term $t = f(f(a,a),f(a,a))$ in the TRS

$$\begin{aligned} a &\to b \\ f(x,y) &\to g(x,x) \end{aligned}$$

An outermost development of t^* witnessing that $\mu(\phi) = 7$ is:

$$\begin{aligned} &\underline{f}(f(a,a),f(a,a)) \to g(\underline{f}(a,a),f(a,a)) \to g(g(\underline{a},a),f(a,a)) \to \\ &g(g(b,\underline{a}),f(a,a)) \to g(g(b,b),\underline{f}(a,a)) \to g(g(b,b),g(\underline{a},a)) \to \\ &g(g(b,b),g(b,\underline{a})) \to g(g(b,b),g(b,b)) \end{aligned}$$

(The length of an *arbitrary* development of t^* is 13, see also Remark 4.28.)

Proposition 4.27 *Let ϕ be a multi-step eliminating a co-initial parallel step ψ. Then $\mu(\phi) \geq \mu(\phi/\psi)$ and if ψ moreover is a head step, then $\mu(\phi) > \mu(\phi/\psi)$.*

Proof We show that, for any multi-step ϕ with ψ/ϕ empty, for any outermost development of ϕ/ψ starting from s, we may construct an outermost development of ϕ starting from t, of at least the same length.

If ϕ is the empty multi-step, there is nothing to prove.

Otherwise, let ψ' be outermost in ϕ/ψ and suppose ψ' is a residual of ϕ'

in ϕ. First note that if the redex-pattern ϕ_1 is above the redex-pattern ϕ_2 in ϕ, then for any residual of ϕ_2, there exists some residual of ϕ_1 above it *unless* some redex-pattern of ψ coincides with ϕ_1. Therefore, since ψ' was assumed to be an outermost (in ψ) residual of ϕ':

- Either ϕ' is outermost (in ϕ) as well. Since we assumed that ψ/ϕ is empty, there cannot be a redex in ψ above ϕ'. Therefore, ψ' is the unique residual of ϕ', and after performing the step ϕ', we may repeat the construction for ϕ/ϕ' and ψ/ϕ'.

  To see that repetition is allowed, note that ψ/ϕ' is a parallel reduction from the target of ϕ' to the target of ψ'. Moreover, it is eliminated by ϕ/ϕ', since $(\psi/\phi')/(\phi/\phi') = (\psi/\phi)/(\phi'/\phi) = \emptyset/(\phi'/\phi) = \emptyset$, by the cube law for residual systems (see Remark 2.26).
- Or there is some outermost redex-pattern ϕ'' directly above ϕ' which is both in ϕ and ψ. Then after performing ϕ'', we may repeat the construction for ϕ/ϕ'' and ψ/ϕ''.

  To see that repetition is allowed, note that ψ/ϕ'' is a parallel step from the target of ϕ'' to the target of ψ, i.e. the source of ψ'. Moreover it is seen to be eliminated by ϕ/ϕ'' as in the previous item.

$\square$

Remark 4.28 Proposition 4.27 would fail if we were to take erased steps into account, e.g. by taking the length of an *arbitrary* development as measure of a multi-step. For instance, consider the multi-step $\phi{:}\underline{e}(d(\underline{a})) \multimap\!\!\!\rightarrow c$ and the co-initial step $\psi{:}e(\underline{d}(a)) \multimap\!\!\!\rightarrow e(f(a,a))$ in the orthogonal TRS with rules

$$\begin{aligned} a &\to b \\ e(x) &\to c \\ d(x) &\to f(x,x) \end{aligned}$$

Then the length of a maximal development of ϕ is 2 which is *less* than the length 3 of a maximal development of its residual $\phi/\psi{:}\underline{e}(f(\underline{a},\underline{a})) \multimap\!\!\!\rightarrow c$ after ψ! (Note that $\mu(\phi) = 1 = \mu(\phi/\psi)$.)

Lemma 4.29 (Non-Head) *If in an orthogonal TRS, a parallel head cycle σ is co-initial to a multi-step ϕ with $\phi/\sigma = \phi$, then ϕ does not eliminate σ.*

Proof Suppose (the successive residuals of) ϕ would eliminate all parallel steps of σ. Then, by repeated application of Proposition 4.27, $\mu(\phi) > \mu(\phi/\sigma) = \mu(\phi)$, since σ was assumed to contain at least one head step. $\square$

These lemmas allow for an alternative proof to Theorem 3.17.

Proof [of Theorem 3.17] For a proof by contradiction, assume there would exist a non-empty parallel cycle. Select among the parallel cycles of minimal length a parallel head cycle, say σ starting with head step ψ from t. For any multi-step ϕ co-initial to σ, Lemma 4.62, yields an extension $\bar{\phi}$ of ϕ. By the Head Lemma $\psi/\bar{\phi}$ is either empty or a head step again. But the former cannot hold, since then $\sigma/\bar{\phi}$ would be a parallel cycle shorter than σ, which by the Non-Head-Cluster lemma is non-empty, contradicting minimality of the length of σ. Therefore, $\sigma/\bar{\phi}$ is a parallel head cycle again. Now by the weak head normalisingness assumption there is a reduction τ from t to head normal form s. Hence, by Lemma 2.9 and the above, we may assume that $\tau/\sigma = \tau$ and that σ/τ is a parallel head cycle. But this would contradict s being in head normal form. $\square$

4.2.2 Preservation of parallel head cycles in weakly orthogonal TRSs

We extend the results for orthogonal TRSs of the previous section to weakly orthogonal TRSs. In particular, we prove analogues of the Head and Non-Head Lemmas.

Our first observation is that the Head Lemma does *not* hold for weakly orthogonal TRSs, as witnessed by the following example.

Example 4.30 Consider the term $t = g(f(f(a)))$ in the TRS with rules

$$\begin{aligned} g(f(f(x))) &\to g(f(x)) \\ f(x) &\to x \end{aligned}$$

Since both critical pairs are trivial, this TRS is weakly orthogonal. Consider the weakly orthogonal projection of the head step $\phi{:}\underline{g(f(f}(a))) \to g(f(a))$ from the term t, over the co-initial step $\psi{:}g(\underline{f}(f(x))) \to g(f(a))$. By the definition (Lemma 2.24) of the weakly orthogonal projection, we first have to canonise both steps: $\iota(\phi){:}g(f(\underline{f}(a))) \to g(f(a))$ and $\iota(\psi) = \psi$. Using this, we compute $\phi /\!\!/ \psi{:}g(\underline{f}(a)) \to g(a)$ and see that the residual of the head step ϕ after ψ is *not* a head step. Stronger, $g(f(a))$ does not even have a head redex!

The problem in the example is that the redex-pattern contracted by ψ is *entirely above* the redex-pattern contracted by the *canonisation* of the head step ϕ. We will show that for given head step ϕ and cycle-preserving multi-step ψ we may always create the property that there are no redex-patterns in ψ which *offend* ϕ in this way, i.e. there are no redex-patterns contracted in ψ which are entirely above the canonical step representing ϕ. This will suffice to regain the Head Lemma. The proof is presented in the next three paragraphs, corresponding to the following three proof steps. First, we perform a careful analysis of *clusters*, which can be thought of as partially characterising the residuals of redex-patterns in a weakly orthogonal TRS. Next, we prove an *externality* [17, Def. 9.2.31] result for head clusters, i.e. clusters containing a head redex: all redex-patterns in a residual of a head cluster descend linearly until contracted (if at all). Finally, we use externality to create the desired property, by *eliminating* from the head cluster of ψ all redex-patterns offending ϕ, one by one.

Clusters are coverable by chains

We perform a careful analysis of so-called clusters, which can be thought of as partially characterising the residuals of redex-patterns in a weakly orthogonal TRS. In particular, we show that clusters must be *coverable* by *chains*, in a sense to be defined below, in (hypothetical) TRSs which are weakly head normalising, weakly orthogonal, and cyclic. This will enable us, in the subsequent paragraph, to show that the residual of a head cluster is a *single* cluster again. To prove that clusters are coverable by chains, we prove that several shapes are *forbidden*, i.e. not possible in such TRSs. For convenience, we assume in this paragraph

$\mathcal{R}$ is an arbitrary weakly head normalising, weakly orthogonal, cyclic TRS.

Definition 4.31 A *cluster* is a non-empty set of redex-patterns which forms a connected component w.r.t. the has-overlap-with relation. A *head* cluster is a cluster containing a head redex-pattern (see Figure 12).

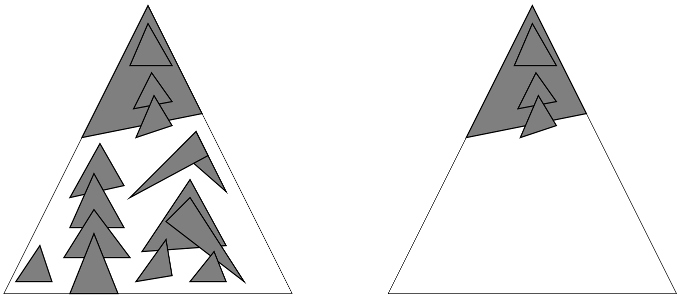

Figure 12. Clusters and head-cluster for the term t of Figure 6.

In other words, two redex-patterns belong to the same cluster if they are related by the reflexive–transitive closure of the (symmetric) has-overlap-with relation, i.e. if there is a sequence of redex-patterns having overlap, which connects them. We often identify a cluster with the set of positions of all its redex-patterns.

Example 4.32 • The term $P(\underline{P(S(P(}S(S(\underline{S(P(}x))))))))$ has the two clusters indicated by the underlining, with respect to the TRS of Example 2.17,

- The whole term $por(\top, \top)$ is a cluster in the TRS with rules

$$\begin{aligned} por(\top, y) &\to por(\top, y) \\ por(x, \top) &\to por(x, \top) \\ \top &\to \top \end{aligned}$$

  the leftmost $\top$-redex is overlapped by the $por(\top, y)$-redex which has overlap with the $por(x, \top)$-redex which overlaps the rightmost $\top$-redex.

Remark 4.33 This notion of cluster generalises that of [6] (cf. Remark 4.49).

Proposition 4.34 *If ϕ, ψ are redex-patterns belonging to the same cluster, then they have the same sources and targets.*

Proof By induction on the length of the sequence of redex-patterns having overlap which connects ϕ to ψ, each time using weak orthogonality. □

Our first impossibility result is that there can be no looping clusters.

Definition 4.35 A cluster is *looping* if it admits a loop (page 4).

Proposition 4.36 (Loop) *$\mathcal{R}$ is* non-looping, *i.e. there are no looping clusters.*

Proof For a proof by contradiction, suppose some looping cluster C would exist in some term t. Clearly the cluster C itself then constitutes a term embedded in t. By the assumption that C is a looping cluster and Proposition 4.34, C only allows looping steps. Since C does allow a head step, this would contradict the assumed weak head normalisation of $\mathcal{R}$. □

Therefore, in the rest of this paragraph we assume $\mathcal{R}$ *is non-looping.* Next, we show that clusters cannot 'fork', they are 'straight lines'.

Definition 4.37 A cluster is a Y-cluster is if it contains parallel redex-patterns. Otherwise, it is an I-cluster (see Figure 13).

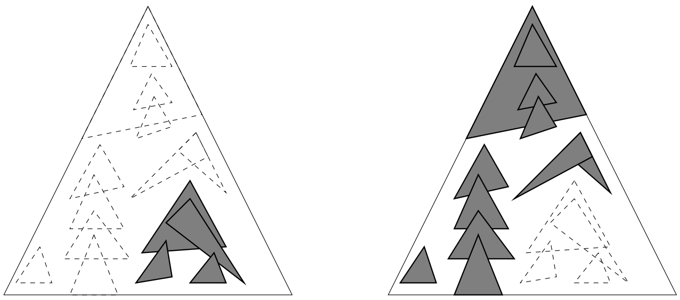

Figure 13. Y-cluster and I-clusters of the term t of Figure 6.

The first item of Example 4.32 exemplifies I-clusters. Since the two $\perp$-redex-patterns in the second item are parallel and belong to the same cluster, the cluster is a Y-cluster.

Remark 4.38 The notion of Y-cluster is a generalisation of the notion of Takahashi-configuration as appearing in [17, Fig. p.12]. The generalisation is proper in the sense that there are Y-clusters which do not contain a Takahashi-configuration. (Hint: Construct a term as in the Y-cluster in Figure 13.)

Proposition 4.39 (Y) *$\mathcal{R}$ is* Y-free, *i.e. there are no Y-clusters.*

Proof Let ϕ, ψ be steps contracting two parallel redexes in a Y-cluster. By Proposition 4.34, ϕ and ψ have the same targets. Note that the only way in which contracting two disjoint redex-patterns can yield the same result is when they are in fact looping [17, Proposition 9.3.5], from which we conclude by Proposition 4.36. □

Therefore, in the rest of this paragraph we assume $\mathcal{R}$ *is* Y-*free*. Since the second item of Example 4.32 displays a Y-cluster, the TRS is not weakly head normalising and therefore not weakly normalising either. Both the TRS in the first item and the $\lambda\beta\eta$-calculus are Y-free.

Corollary 4.40 *For any cluster in a term, there is a path from the root to a leaf, in the term seen as a tree, through all head positions of all redex-patterns in the cluster.*

Proof Because clusters are I-clusters, the set of head positions of redex-patterns is linearly ordered, i.e. is a subset of some path from the root to the leaves. □

In the rest of this paragraph we fix some such path, which we call π.

Finally, we show that clusters may be assumed to be free from 'bridges', i.e. from redex-patterns which have overlap with two nested redex-patterns 'bridging' (hence having overlap with) a redex-pattern in-between.

Definition 4.41 Let t be a term containing three redex-patterns ϕ_1, ϕ_2, ϕ_3, such that ϕ_i is entirely above ϕ_{i+1}. Then a redex-pattern ψ is said to *bridge* ϕ_2 *from* ϕ_1 *to* ϕ_3 if it has overlap with each ϕ_i (see Figure 14 left).

Example 4.42 In the term $\underline{h(g}(\overline{\underline{f}(\underline{f}(a)})))$, the over-lined redex-pattern bridges the middle f-redex-pattern, for the weakly orthogonal TRS with rules:

$$\begin{aligned} h(g(x)) &\to h(g(f(a))) \\ g(f(f(a))) &\to g(f(a)) \\ f(x) &\to x \end{aligned}$$

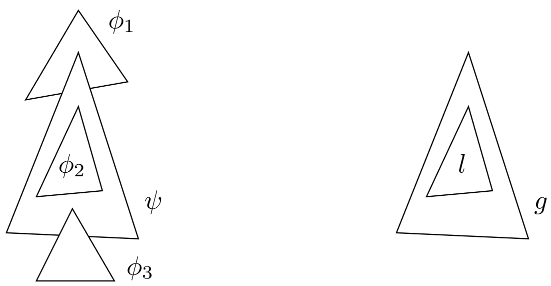

Figure 14. ψ bridges ϕ_2 from ϕ_1 to ϕ_3, and l is embedded in g

Note that the second rule in the example is more specific than the third rule, in the sense that the left-hand side of the latter is embeddable into the left-hand side of the former (see Figure 14 right). We will show that this holds in general for bridges, from which absence of bridges follows since we may assume absence of embeddings, as we show now.

Proposition 4.43 (Embedding) *We may assume that* $\mathcal{R}$ *is* non-embedding, *i.e. that left-hand sides of distinct* $\mathcal{R}$*-rules are mutually non-embeddable.*

Proof Let $\mathcal{R}'$ be obtained from $\mathcal{R}$ by omitting any rule whose left-hand side is non-minimal with respect to the embedding relation. We verify that $\mathcal{R}'$ has all properties $\mathcal{R}$ has. First note that the rewrite relations $\to_{\mathcal{R}}$ and $\to_{\mathcal{R}'}$ coincide. This holds, since if $t \to_{\mathcal{R}} s$ for some rule $g \to d$ in $\mathcal{R}$ such that g is not minimal w.r.t. the embedding relation, there is by well-foundedness of the embedding relation, a rule $l \to r$ in $\mathcal{R}'$ such that l is embeddable into g. Hence, by weak orthogonality $t \to_{\mathcal{R}'} s$.

- Since $\mathcal{R}$ is weakly head normalising and all $\mathcal{R}$-head normal forms are $\mathcal{R}'$-head normal forms, $\mathcal{R}'$ is weakly head normalising.
- $\mathcal{R}'$ is weakly orthogonal, as sub-TRS of the weakly orthogonal TRS $\mathcal{R}$.
- Since $\to_{\mathcal{R}}$ and $\to_{\mathcal{R}'}$ coincide, either is cyclic if the other is.

□

Therefore, in the rest of this paragraph we assume $\mathcal{R}$ *is non-embedding.*

Remark 4.44 Proposition 4.43 invalidates the TRS used in Example 4.30, since its first rule embeds the second one. However, head redex-patterns need not be preserved in non-embedding TRSs either, as witnessed by the term $g(h(f(f(a)), a))$ in the non-embedding weakly orthogonal TRS with rules

$$\begin{aligned} g(h(f(f(x)), y)) &\to g(h(f(x), y)) \\ h(f(x), a) &\to h(x, a) \\ f(a) &\to a \end{aligned}$$

as can be seen by proceeding in the same way as in Example 4.30.

A redex-pattern ϕ may be mapped to the set $[\phi]$ of its positions on π.

Lemma 4.45 *Let ϕ, ψ be redex-patterns in the same cluster. Then*

- *$[\phi]$ is a convex subset of π containing the head position of ϕ.*
- *ϕ has overlap with ψ iff $[\phi]$ has a non-empty intersection with $[\psi]$, and*
- *ϕ is entirely above ψ iff every position in $[\phi]$ is strictly above every position in $[\psi]$.*

Proof

- By convexity of redex-patterns [17, Sect. 8.6.2] and since the position of the head of ϕ is on π by assumption.
- The if-direction being trivial, we prove the only-if-direction. Suppose o is a position both in ϕ and ψ. Since the head positions p, q of ϕ, ψ are linearly ordered by the prefix order, we may assume w.l.o.g. that p is above q (not necessarily strictly so). By the previous item applied to p and o in ϕ, q then belongs to ϕ as well. The result follows since q certainly is on π.
- By the previous items.

□

By the first item of the lemma, the set of positions $[\phi]$ associated to the redex pattern ϕ is an *interval* in the sense of [2]. By its second and third item, we may replace abstract reasoning only involving the notions 'overlap with' or 'entirely above of' for redex-patterns ϕ and ψ, by reasoning on their associated intervals $[\phi]$ and $[\psi]$. Therefore, we will below often identity a redex-pattern with its associated interval. We say that the interval ϕ *nests* the interval ψ, $\phi \prec \psi$, if every position in ϕ is strictly above every position in ψ (using the terminology of [2] this is expressed as: ϕ is *before* or *meets* ψ). If either interval nests the other, we say they are disjoint $\phi \perp \psi$. Otherwise they are said to *have overlap* $\phi \not\perp \psi$.

Proposition 4.46 (Bridge) *$\mathcal{R}$ is* bridge-free, *i.e. clusters don't have bridges.*

Proof For a proof by contradiction, let C be a cluster containing three redex-patterns ϕ_1, ϕ_2, ϕ_3, such that ϕ_i is entirely above ϕ_{i+1}, and let ψ bridge ϕ_2 from ϕ_1 to ϕ_3. Let the respective rules be $l_i \to r_i$ for $i \in \{1,2,3\}$ and $g \to d$ and let ϕ_{i+1} occur in the argument for variable x_i in the left-hand side l_i. Without loss of generality we may assume that all positions of C belong to (at least one) of the ϕ_i or ψ.

We distinguish cases on the different possible shapes of ϕ_2.

- That ϕ_2 could be embedded into ψ, would contradict Proposition 4.43.
- Otherwise, there is a position p of a symbol which is in ϕ_2 but not in ψ.

  If p were strictly above ψ, then by ϕ_1 being entirely above ϕ_2, ϕ_1 would be entirely above ψ as well, but then $[\phi_1]$ and $[\psi]$ would not have overlap, so by Lemma 4.45 neither would ϕ_1 and ψ. Symmetrically, if p would be strictly below ψ but comparable to some position of ϕ_3, then since ϕ_3 is entirely below ϕ_2, a contradiction with ϕ_3 and ψ having overlap is obtained.

  Hence p must occur in some argument of g, at a position incomparable to (the positions of) ϕ_3. We claim that therefore x_1 must occur exactly once in r_1, and at the same position as in l_1. To see this, prune the sub-term at position p, replacing it by a fresh variable, say y, yielding the term D.

  To prove the claim, first note that D still contains the redex-patterns ϕ_1, ψ and ϕ_3, with ψ having overlap with both the others. Hence contracting any of them in D always yields the same result. Next note that contracting ϕ_3 does not change the position of y, since it is incomparable to it. Hence contracting ϕ_1 does not change it either. Thus the positions of x_1 must be exactly the same in the left- and right-hand sides of $l_1 \to r_1$. Since the rule is left-linear, the claim follows.

  Now, on the hand, contracting ϕ_1 does not change the part of C *below* x_1, by the claim. On the other hand contracting ϕ_3 does not change the parts of C *above and disjoint from* x_1. Since both steps must have the same result being part of the same cluster, we therefore conclude that all steps in the cluster C are looping, contradicting Proposition 4.36.

□

The Bridge Proposition will be essential for the residual of a head cluster not to fall apart into several clusters, but to remain one. The reason for this is that although the redex-patterns in a cluster may be overlapping one another in complex ways, there always is a simple structure *covering* all of them: a *chain* (see Figure 15).

Definition 4.47 Define for sets of redex-patterns Φ, Ψ their difference $\Phi - \Psi$ as the set obtained by removing from Φ all redex-patterns which have overlap with some redex-pattern in Ψ. We say that Φ is *covered* by Ψ, if $\Phi - \Psi$ is the empty set. A *chain* is a sequence of (intervals of) redex-patterns such that each is overlapped by[10] its successor, but doesn't have overlap with *its* successor.

The redex-patterns (intervals) of a chain will be called its *links*.

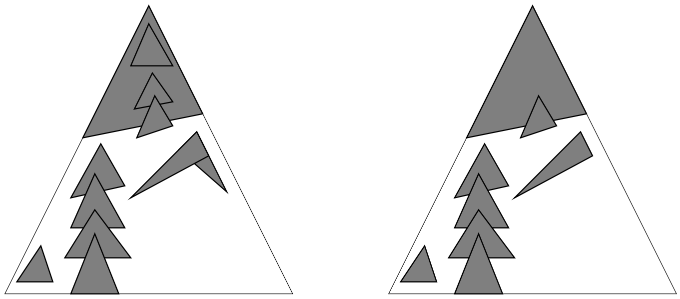

Figure 15. Some I-clusters and their covering chains

Example 4.48 Consider the term t of Example 2.20 in the weakly orthogonal TRS of Example 2.17. It is a head cluster which is covered by the sequence of *all* its redex-patterns. More generally, taking *all* redex-patterns in a cluster yields a chain, in either that TRS or in the other rewriting system of Example 2.17, the $\lambda\beta\eta$-calculus, as one easily verifies. That is, the situation is as for the bottom–middle I-cluster in Figure 15.

Remark 4.49 Our definition of chain is in concordance with the definitions of *chain of* λs as found in [5] and of *chain* as in [8].

We construct a chain covering C by a greedy algorithm: Starting from the bottom of C, always adjoin the 'greatest possible' link. We first present the algorithm, after which we prove in the subsequent Chain Lemma that it indeed generates a chain covering C.

Definition 4.50 The *greedy chain* $\circledcirc_C$ of a cluster C is constructed by:

- As first link choose a redex-pattern of C such that the lowest position of C on π is in it, and among those one with the highest head position.

[10] Here we use overlapped-by in the sense of [2]: ϕ is *overlapped by* ψ if they have overlap, i.e. a non-empty intersection, there is a position of ϕ strictly below ψ, and, vice versa, there is a position of ψ strictly above ϕ.

- Search for the redex-pattern in C with the highest head position and such that adjoining it yields a chain. If it exists, adjoin it as next link and repeat.

Lemma 4.51 (Chain) *$\circledcirc_C$ is a chain covering C.*

Proof That the greedy chain is a chain holds trivially by construction.

Suppose $\circledcirc_C$ would not cover C. Consider a redex-pattern ϕ in C which does not have overlap with any redex-pattern of the greedy chain. Without loss of generality we may assume that it extends maximally downward among those, and thus the sequence witnessing that ϕ and the final link of $\circledcirc_C$ are connected (are being part of the same cluster) consists of a single link, say ψ. Now note that ψ has overlap with $\circledcirc_C$, but that its head position is above all heads of greedy redex-patterns. Consider the first greedy redex-pattern χ with which it has overlap. If χ is the first link of $\circledcirc_C$ and the lowest position of C is in ψ, then ψ should have been chosen instead of χ. Otherwise, ψ should have been chosen as the successor of χ. Contradiction. $\square$

We now use the greedy chain in our proof of the main result of this paragraph. It expresses that after removing a suffix, i.e. a downward closed subset of the maximal set of redex-patterns of a cluster, and all redex-patterns having overlap with it from the cluster, the remaining redex-patterns still form a 'cluster'. This is based on the fact that removing a suffix from a chain yields a chain again, whereas removing an arbitrary set of links could obviously split the chain into a number of disconnected parts.

Theorem 4.52 (Suffix) *For any suffix Φ of C^*, $C^* - \Phi$ has a covering chain in $C - \Phi$.*

Proof Since in case $C^* - \Phi$ is empty or a singleton the statement holds trivially, we assume the set has at least two elements. We claim that $\circledcirc_C - \Phi$, which clearly is in $C - \Phi$, is a chain covering $C^* - \Phi$. To that end it suffices to show that the lowest element of $C^* - \Phi$, which we denote by ϕ_2, is covered by some link of $\circledcirc_C - \Phi$. (This suffices by downward closedness of Φ and $\circledcirc_C$ being a chain, so that each link is overlapped by its successor, meaning that its successors are above it.) We claim that the last link (generated in the construction) of $\circledcirc_C$ which had overlap with ϕ_2, which we denote by ψ, is not removed when taking the difference $\circledcirc_C - \Phi$. For suppose ψ would have been removed, i.e. suppose that it had overlap with some redex ϕ_3 in Φ. By Φ being a downward closed subset of C^*, ϕ_2 nests ϕ_3. We distinguish cases on whether there is a successor to ψ in $\circledcirc_C$ or not, in both cases finding a redex-pattern ϕ_1 such that ψ bridges ϕ_2 from ϕ_1 to ϕ_2, yielding a contradiction with the Bridge Proposition.

Suppose there is a successor to ψ in $\circledcirc_C$. Then we take that to be ϕ_1. By $\circledcirc_C$ being a chain, ϕ_1 must have overlap with ψ, and by the choice of ψ being the last link of $\circledcirc_C$ having overlap with ϕ_2, ϕ_1 nests ϕ_2. Hence the preconditions of the Bridge Proposition are satisfied.

If there is no successor to ψ in $\circledcirc_C$, then let ϕ_1 be a redex-pattern in $C^* - \Phi$ distinct from ϕ_2, which exists by assumption. By ϕ_2 being lowest, it is nested by ϕ_1. Since $\circledcirc_C$ covers C with ψ as last link, which has overlap with ϕ_2, ψ must have overlap with ϕ_1 as well, as the latter is above ϕ_2, the preconditions of the Bridge Proposition are satisfied again, and we are done. $\Box$

Remark 4.53 Both the Chain Lemma and the Suffix Theorem are abstract in the sense that they only employ manipulation with intervals. Hence these results hold for finite clusters in interval algebra [2]. For instance, on intervals a *cluster* is a (finite) set of intervals such that their union is a single interval, a *cover* is a subset having the same union, a *chain* is as expected, and a *bridge* is a cluster consisting of four intervals three of which are pair-wise non-overlapping. Then the Chain Lemma expresses that every cluster without bridges has a chain as cover.

Head clusters are external

Here we prove that this implies that any head cluster C is *external* [17, Def. 9.2.31], in the sense that its residuals cannot be nested by redex-patterns not belonging to the cluster, along a reduction σ. For this to hold, we use the fact from the previous paragraph that any cluster may be covered by a chain. In order for this property to be preserved by projection, i.e. for the residual of a chain to be a chain again, we require that the steps of the reduction σ are *down* steps, causing only links at the tail of the chain to be removed when taking residuals. Externality of C and steps of σ being down steps, together will yield the property that all redex-patterns in C have unique residuals, until possibly contracted. Uniqueness will be the basis for their elimination in the subsequent paragraph.

We define the residual of a cluster C via its maximal step ϕ^*, and adapt the notion of externality of steps as presented in [17, Def. 9.2.31] accordingly.

Definition 4.54 The set of clusters *generated* by a subset P of the positions of a term, denoted by $\overline{P}$, is the least set of clusters which contains P. The *residual* of a cluster C after a multi-step ϕ is defined by $\overline{C^* /\!\!/ \phi}$. A cluster C in a term t is *external to* a reduction σ from t if the residuals of C are not nested by other (not necessarily contracted) redex-patterns in the course of σ.

Note that even if a cluster is external, so there are no redex-patterns *outside* its residuals which could duplicate its residuals, duplication could be caused by redex-patterns *within* the cluster itself, invalidating the desired uniqueness.

Example 4.55 Consider the cluster $f(d(x, y))$ in the weakly orthogonal TRS

$$\begin{aligned} f(x) &\to f(d(x, x)) \\ f(d(x, y)) &\to f(d(d(x, y), d(x, y))) \end{aligned}$$

$$d(x, y) \to d(d(x, y), d(x, y))$$

Contracting the f-redex leaves *two* residuals of the d-redex in the same cluster. [11]

Note that the f-redex-pattern is *non-down* in the sense that there is a non-contracted redex-pattern which is below it and in the same cluster, namely the d-redex-pattern. We show that any parallel reduction can trivially be transformed into one performing only down-steps, of the same length. (In the case of the example, the f-redex will be transformed into the d-redex.)

Definition 4.56 A *down* step from t, is a sub-step ϕ of t^* such that, for any redex-pattern ψ in ϕ, all redex-patterns of t^* below and in the same cluster as ψ, are in ϕ as well (see Figure 16).

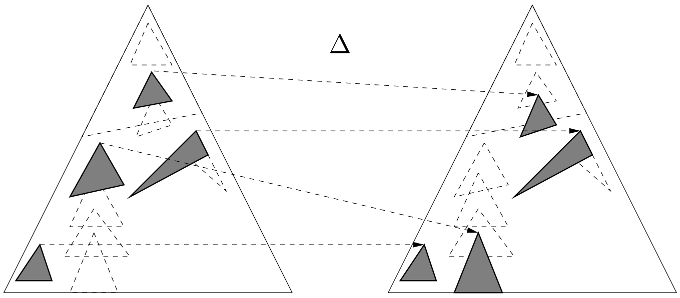

Figure 16. A multi-step χ and its corresponding down step $\Delta(\chi)$

Lemma 4.57 *For any multi-step* $\phi{:}t \multimap s$ *there exists a down step* $\Delta(\phi){:}t \multimap s$*. Moreover,* ϕ *and* $\Delta(\phi)$ *contract the same number of redex-patterns,* $\Delta(\phi) \subseteq \phi^*$*, and if* ϕ *is a parallel step, then* $\Delta(\phi)$ *is one as well.*

Proof See Figure 16. Formally, one reasons by induction on the number of non-down redex-patterns contracted in ϕ. In case this number is 0, we are done and may simply set $\Delta(\phi) = \phi$. Otherwise, the number is $n + 1$ for some n, and we may remove one innermost non-down redex-pattern and replace it by a down redex-pattern in the same cluster and below it, but not in ϕ. Say this yields ϕ'. By Proposition 4.34, ϕ and ϕ' have the same sources and targets and they obviously contract the same number of redex-patterns. However, in ϕ' there is one more redex down, and only n redexes to go. Therefore, we conclude by the induction hypothesis, setting $\Delta(\phi) = \Delta(\phi')$.

By Y-freeness, the construction is deterministic, except in case a choice exists between two redex-patterns at the same position, having overlapping heads, at the same maximal depth. In such cases, we make the same choice as in the construction of the maximal multi-step t^*, yielding $\Delta(\phi) \subseteq t^*$.

[11] Both the first and third rules are embeddable into the second rule. As in Example 4.44 this invalidates the example (being a counter-example), but this can be overcome likewise, e.g. by providing f and d with extra arguments (which are to be instantiated in the first and third rule, but not so in the second rule).

That this procedure preserve parallel-ness of multi-steps follows, since replacing a redex-pattern by one below it preserves parallel-ness. □

Example 4.58 We illustrate Δ on the same steps $\phi{:}\overline{P(S}(P(S(P(x))))) \multimap\!\!\!\to P(S(P(x)))$ and $\psi{:}\overline{P(S}(\overline{P(S}(P(x))))) \multimap\!\!\!\to P(x)$ and from the same term $t = P(S(P(S(P(x)))))$, as was used in Example 2.22 to illustrate canonisation. Since t constitutes a single cluster, the down step corresponding to ϕ, simply contracts the innermost redex-pattern in t i.e. $\Delta(\phi){:}P(S(P(\underline{S(P}(x))))) \multimap\!\!\!\to P(S(P(x)))$. It is easy to see that $\Delta(\psi) = t^*$. The difference between canonical and down steps, as is well-illustrated by ϕ, is that $\iota(\phi)$ is not necessarily a down step, and, vice versa, $\Delta(\phi)$ does not necessarily have overlap with ϕ. (But both may hold, as illustrated by ψ.)

Taking residuals of redex-patterns in non-nested clusters is particularly simple; either the redex-pattern was not contracted and then it projects onto itself, or else it was contracted and then it leaves no residual. Using this characterisation, we establish some further technical properties which will be useful in the proof of Theorem 4.60. For instance, the final item characterises the situation in Example 4.30 as being completely general.

Lemma 4.59 *Let $\phi{:}t \multimap\!\!\!\to s$ be a canonical multi-step co-initial to a down step ψ. Let ϕ contract redex-patterns in a non-nested cluster C at position p. Then*

(i) $\phi /\!\!/ \psi = \iota(\phi - \psi)$, *and*

(ii) *$C /\!\!/ \psi$ is empty iff $\chi /\!\!/ \psi$ is empty, with χ outermost among C^*.*

(iii) *the head positions of C and $C /\!\!/ \psi$ are comparable (if the latter exists).*

(iv) *If C is not eliminated, but there is no redex-pattern at position o in s, then any redex-pattern ζ at o in t has overlap with exactly two redex-patterns in C^*, the outermost one of which is also a redex-pattern in $C /\!\!/ \psi$.*

Proof We show the items in order. For the first item, note that by definition, by assumption for ϕ, and since down-steps are canonical for ψ, we have $\phi /\!\!/ \psi = \iota(\iota(\phi)/\iota(\psi)) = \iota(\phi/\psi)$. Hence it suffices to verify that

$$\phi/\psi = \phi - \psi$$

holds. This follows by inspection of the four clauses for residuation of multi-steps as presented in [17, Def. 8.7.4]. More precisely, one shows that because the cluster C was assumed to be non-nested and ψ is a down step, the second clause there never can be applied anymore once either the third or the fourth clause has been applied.

For the second item, note that by definition $C /\!\!/ \psi = \overline{C^* /\!\!/ \psi}$ and by assumption $\chi \in C^*$, so $C /\!\!/ \psi$ is non-empty if $\chi /\!\!/ \psi$ is non-empty. To show the other direction, note that by the first item $\chi /\!\!/ \psi = \iota(\chi - \psi)$. Since ψ is a downward-closed sub-step of C^* by assumption, and χ is outermost in C^*, $\chi - \psi$ is only going to be empty if $\psi = C^*$, and the result follows.

To show the third item, assume that p and p' are the positions of the heads of C and $C /\!\!/ \psi$. Since $C /\!\!/ \psi$ is non-empty, we know by the previous item that $\chi /\!\!/ \psi$ is non-empty too, with χ as in that item. Clearly p and p' are above the positions q and q' of χ and $\chi /\!\!/ \psi$, respectively. By the first item, $\chi /\!\!/ \psi = \iota(\chi - \psi)$, so $\chi /\!\!/ \psi = \iota(\chi)$, thus q' is seen to be below q, by inspection of canonisation. Therefore, both p and p' are above q', so they are comparable since positions form a tree.

It remains to prove the fourth and final item. If for some outermost redex-pattern ζ of C, no redex-pattern having overlap with ζ would be contracted by ψ, we would have that ζ is in u as well (use the displayed equality in the first item). Therefore, for any outermost redex-pattern ζ of C, there is at least one redex-pattern ζ' in ψ having overlap with it. By the assumption that C is not eliminated, we have by the second item that the outermost redex-pattern χ in C^* is not eliminated by ψ. Therefore it is distinct from ζ', so by its outermost-ness, χ nests ζ'. We conclude since χ has overlap with ζ, because ζ' does so, and there cannot be other redex-patterns (having overlap with ζ, but not with either χ or ζ'), by the Bridge Proposition. □

Theorem 4.60 (External) *Head clusters are external to any down reduction. Moreover, a redex-pattern in a residual of the head cluster along the reduction, has unique residuals in* its *residual clusters until contracted.*

Proof The proof is by well-founded induction on the length of down reductions from a term t with head cluster C. Hence, assume the statement holds for all down reductions up to a given length. If the length is zero, the statement holds trivially. Otherwise, suppose the reduction is of shape $\sigma \cdot \psi$, with σ a down reduction from t to s and ψ a down step from s to u. By the induction hypothesis, the statement holds for σ, so we may define $D = C /\!\!/ \sigma$ to be the unique residual cluster of C along σ, and it holds that there is no redex-pattern strictly above D. Moreover, redex-patterns in any residual of C have unique residuals in D, until contracted. We show that $E = D /\!\!/ \psi = \overline{D^* /\!\!/ \psi}$ in u has these properties again.

We first show that E consists of *at most one* cluster. As there is no redex-pattern nesting D and since ψ is a down step, Lemma 4.59(i) yields $D^* /\!\!/ \psi = \iota(D^* - \psi)$. Since ψ is a down step, removing it consists in removing a suffix from the chain covering D^*. Therefore, by the Suffix Theorem 4.52, there is a chain in $D - \psi$ covering $D^* - \psi$, i.e. all residual steps belong to the same cluster again. Since the property of being a single cluster is preserved both by canonisation and cluster generation, we conclude that E is a single cluster.

Next we show that E is *non-nested*, i.e. that there is no redex-pattern entirely above it. We present a proof by contradiction based on the idea that if such a redex-pattern would exist, then the reduction σ can be rearranged to witness a bridge, contradicting the Bridge Proposition (see Figure 17). So assume existence of such a redex-pattern ϕ_1 in u. Then E is non-empty and we may consider a position p in ϕ_1, which is at maximal depth on the path

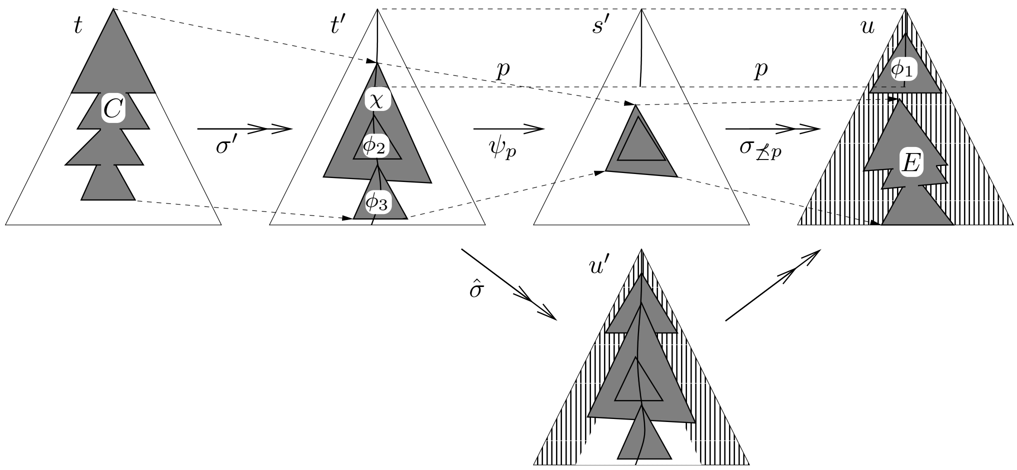

Figure 17. Decomposing $\sigma \cdot \psi$ into $\sigma' \cdot \psi_p \cdot \sigma_{\not\geq p}$ and extracting $\hat{\sigma}$ ending in bridge u'

π from the root of s through the heads of redex-patterns in E. Note that C is a head cluster, so its position is the root which is above p, and that p is in its turn strictly above the residual E of C. Therefore, we can decompose $\sigma \cdot \psi$ uniquely as $\sigma' \cdot \psi_p \cdot \sigma_{\not\geq p}$, for some down reduction σ' from t to t', some down step ψ_p from t' to s', and a down reduction $\sigma_{\not\geq p}$ from s' to u which is of maximal length such that all residuals of C along it are strictly below p. As the head of $C' = C /\!/ \sigma'$ is not strictly below p per construction, it must be above p by Lemma 4.59(iii). On the other hand, also per construction, the head of $D' = C' /\!/ \psi_p$ is strictly below p (this is witnessed by the dashed 'trace lines' crossing one another for ψ_p in the figure). Therefore the head redex χ of C' is eliminated by ψ_p, but C' itself is not, since it has the non-empty E as residual. By Lemma 4.59(iv) we obtain two nested redexes, say ϕ_2 nests ϕ_3, in C' and having overlap with χ.[12] Moreover, it yields that ϕ_2 is a redex-pattern in D' so it is strictly below p and therefore ϕ_3 is so as well.

Summarising the above, we have found three redex-patterns ϕ_1, ϕ_2 and ϕ_3 such that ϕ_i is entirely above ϕ_{i+1}. Moreover the step χ has overlap with all three of them. In order to derive a contradiction with the Bridge Proposition, it remains to show that they may be assumed to be redex-patterns in the *same* term; presently ϕ_1 is a redex-pattern in u while the others are redex-patterns in t' (see the figure)! Thereto we show that, without affecting the latters, t' can be reduced further to some term u' containing ϕ_1 as well. Note that the reduction $\psi_p \cdot \sigma_{\not\geq p}$ does not contract redex-patterns above p, since ψ_p is a down step which doesn't contract the redex-pattern ϕ_2 below p, and since $\sigma_{\not\geq p}$ doesn't contract redex-patterns above p per its construction and the induction hypothesis. Therefore, we can decompose the reduction $\psi_p \cdot \sigma_{\not\geq p}$. into reductions σ_o which only contract redex-patterns below o, for every position o directly below some position on the path from the root to p, but not on the

[12] The present proof only employs that there are *at least* two nested redexes. It might be the case that it can be simplified using the stronger fact that there are *exactly* two of them.

path itself.

Now consider extracting from $\psi_p \cdot \sigma_{\not\geq p}$ those σ_o for which o is not (strictly) below p, performing these first, say as $\hat{\sigma}$ ending in u' and the others afterwards (see the figure). Then, per construction u' is the same as u for all positions not strictly below p. This means that u' contains the redex-pattern ϕ_1, since p was assumed a minimal position in ϕ_1. Moreover, since the positions below p are unaffected, neither are the the redex-patterns ϕ_2 and ϕ_3. We claim that also the redex-pattern χ is still present in u', which will yield the desired contradiction with the Bridge Proposition. To prove the claim, note that any redex-pattern χ' contracted in $\hat{\sigma}$ is not above p, so to eliminate the redex-pattern χ which *is* above p, the first such χ' must have overlap with χ at a position disjoint from p. This implies that the term at that moment would contain the Y-cluster consisting of χ, χ' and ϕ_2 contradicting the Y-Proposition. This concludes our verification that E is non-nested.

Finally, we check that any redex-pattern ϕ in D has a *unique* residual, unless contracted in ψ. As $\phi /\!\!/ \psi = \iota(\phi) /\!\!/ \psi = \iota(\iota(\phi) - \psi)$, by Lemma 4.59(i), we conclude since canonisation always yields a unique residual, and set-difference does so as well unless the redex-pattern is contracted in ψ. □

Eliminating offending redex-patterns from head-clusters

We use externality, as established in the preceding paragraph, to create the property that when projecting a head cycle over a multi-step, the latter contracts either the (representative of the) head-redex-pattern or no redex-pattern in the head-cluster at all. This is achieved by *eliminating* all *offending* redex-patterns from the multi-step, one by one. To express the results we will use the notion of a parallel head-cluster cycle.

Definition 4.61 A *head-cluster* step is a step contracting a (single) redex-pattern in the head cluster (which is therefore implicitly assumed to exist). A parallel head-cluster cycle is a parallel cycle on t, the first step of which is a head-cluster step.

Lemma 4.62 (Elimination) *Let ϕ be a multi-step co-initial to a parallel head cycle σ on t, which starts with the head step ψ. Then there exist a multi-step $\hat{\phi}$ extending ϕ, and co-initial to a parallel head-cluster cycle $\hat{\sigma}$ on t, starting with the head-cluster step $\hat{\psi}$, such that $\hat{\phi} /\!\!/ \hat{\sigma} = \hat{\phi}$ and either $\hat{\psi} \in \hat{\phi}$ or $\hat{\phi}$ contracts no redex-patterns in the head cluster C of t at all.*

Proof By Lemma 4.57, $\Delta(\sigma)$ is a parallel cycle again, which starts with the single step $\Delta(\psi)$, which contracts the innermost redex-pattern of C^* by definition of Δ. By Theorem 4.21 , there exists a multi-step $\bar{\phi}$ extending ϕ such that $\bar{\phi} /\!\!/ \Delta(\sigma) = \bar{\phi}$, hence $\Delta(\sigma) /\!\!/ \bar{\phi}$ is a cycle again.

If $\Delta(\psi) \in \bar{\phi}$, we may simply set $\hat{\phi} = \bar{\phi}$ and $\hat{\sigma} = \Delta(\sigma)$. Hence, suppose $\Delta(\psi)$ is not in $\bar{\phi}$. We construct $\hat{\phi}$ from $\bar{\phi}$, by eliminating, one by one, the redex-patterns contracted by $\bar{\phi}$ which are also in C. Let this set of *offending* redex-patterns be denoted by $\bar{\phi}_C$ and let χ be an innermost offending redex-pattern in it. By the External Theorem 4.60 χ has a unique residual along $\Delta(\sigma)$, until possibly contracted.

Suppose χ is not eliminated, i.e. has a residual after $\Delta(\sigma)$. We claim that setting $\hat{\phi} = \bar{\phi} - \bar{\phi}_C$ and $\hat{\sigma} = \Delta(\sigma)$, then works. To prove the claim, note that the situation must be as in Figure 18, that is, χ has itself as unique residual after $\Delta(\sigma)$. This holds since by the External Theorem χ has a unique

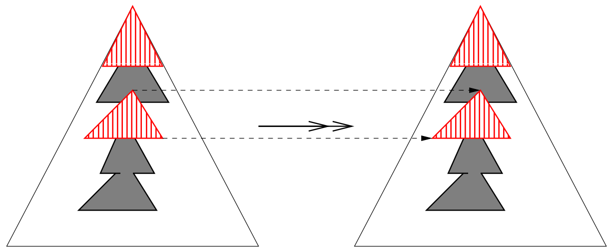

Figure 18. Head cluster with non-eliminated redex-patterns striped/in red

residual after $\Delta(\sigma)$ in the residual of the cluster C, which is non-nested. But since $\Delta(\sigma)$ is a cycle, that non-nested cluster must be C itself, so $\bar{\phi}_C /\!\!/ \Delta(\sigma) \subseteq C /\!\!/ \Delta(\sigma) = C$. Moreover, the residual of a redex-pattern in the head cluster is always below its origin as one easily checks (the tracing arrows in Figure 7 don't point upward), so the residual of χ after $\Delta(\sigma)$ is in $\bar{\phi}_C$ and below χ. But χ was chosen innermost among the redex-patterns in $\bar{\phi}_C$, so χ must in fact be its own residual, as displayed in the figure. To finish the proof of the claim, note that since $\Delta(\sigma)$ consists of down steps, if χ has a unique residual, then all other offending redex-patterns (only one in the figure) have unique residuals as well, since they are entirely above χ by innermost-ness of the latter. By reasoning as for χ, using the assumption $\bar{\phi} /\!\!/ \Delta(\sigma) = \bar{\phi}$, we conclude that in fact, every redex-pattern in $\bar{\phi}_C$ must descend to itself, so $\bar{\phi}_C /\!\!/ \Delta(\sigma) = \bar{\phi}_C$, and hence its complement $\hat{\phi}$ must descend to itself as well: $\hat{\phi} /\!\!/ \Delta(\sigma) = \hat{\phi}$. Since this shows that cyclicity is preserved, it remains to show that $\hat{\phi}$ is an extension of $\bar{\phi}$. At first sight this is not clear, since $\hat{\phi}$ contains *fewer* redex-patterns than $\bar{\phi}$; the offending redex-patterns $\bar{\phi}_C$ have been removed. However, consider replacing in the cycle $\Delta(\sigma)$ its first step $\Delta(\psi)$ by χ. Since both these steps belong to C they have the same sources and targets, hence this replacement gives rise to a cycle again, say σ' on t, Using the assumption that $\Delta(\psi)$ is not in ϕ but is below all redex-patterns in $\bar{\phi}_C$ by it being a down step, we compute

$$\bar{\phi}_C /\!\!/ \chi = \bar{\phi}_C - \chi = (\bar{\phi}_C /\!\!/ \Delta(\psi)) - \chi$$

Since redex-patterns in the same cluster yield the same residual relation on redex-patterns outside the cluster, we also have $\hat{\phi} /\!\!/ \chi = \hat{\phi} /\!\!/ \Delta(\psi)$, so combining both gives $\bar{\phi} /\!\!/ \chi = (\bar{\phi} /\!\!/ \Delta(\psi)) - \chi$ and continuing this along σ' yields

$$\bar{\phi} /\!\!/ \sigma' = (\bar{\phi} /\!\!/ \Delta(\sigma)) - \chi = \bar{\phi} - \chi$$

Hence there is a reduction from the target of $\bar{\phi}$ to the target of the same step where χ has not been contracted. This may be repeated, inside-out, for each offending redex-pattern, eventually yielding a reduction from the target of $\bar{\phi}$ to the target of $\bar{\phi} - \bar{\phi}_C$, i.e. $\hat{\phi}$, finishing the proof of the claim.

Now suppose χ is eliminated, i.e. does not have a residual after $\Delta(\sigma)$. Then we claim that defining $\hat{\phi} = \bar{\phi}$ and taking $\hat{\sigma}$ to be obtained from $\Delta(\sigma)$ by replacing $\Delta(\psi)$ by χ works. For, computing as before, we find

$$\bar{\phi}/\!\!/\chi = (\bar{\phi}/\!\!/\Delta(\psi)) - \chi$$

Continuing this along $\hat{\sigma}$ and using the assumption χ has a unique residual along $\Delta(\sigma)$ until it is eliminated, which it is by assumption, we conclude

$$\bar{\phi}/\!\!/\hat{\sigma} = \bar{\phi}/\!\!/\Delta(\sigma) = \bar{\phi}$$

□

We call the constructed multi-step $\hat{\phi}$ the *head* extension of ϕ and $\hat{\psi}$ the head-cluster step *of* ψ. The hard work having been done, the proofs of the adaptations of the Head and Non-Head Lemmas from the orthogonal to the weakly orthogonal case are trivial.

Lemma 4.63 (Head-Cluster) *Let $\hat{\psi}$ be the head-cluster step co-initial to the head extension $\hat{\phi}$ as in Lemma 4.62. Then $\hat{\psi}/\!\!/\hat{\phi}$ is either empty or a head-cluster step again.*

Proof Immediate from Lemma 4.62. □

To adapt the Non-Head Lemma lengths of (outermost) developments must be invariant under canonisation; they are.

Lemma 4.64 *The lengths of outermost developments*[13] *of a multi-step ϕ and its canonisation $\iota(\phi)$ are the same.*

Proof We may transform a multi-step ϕ into its canonisation $\iota(\phi)$, by replacing inside-out every redex-pattern ψ in ϕ by its canonisation $\iota(\psi)$. Hence it suffices to show that every such single replacement does not change the length of the outermost developments. To that end, note that the unification of the redex-patterns ψ and $\iota(\psi)$ (which have overlap) is disjoint from all other redex-patterns, and can be treated as a single redex-pattern, contracted in either way. □

Lemma 4.65 (Non-Head-Cluster) *Let $\hat{\sigma}$ be the parallel head-cluster cycle co-initial to the head extension $\hat{\phi}$ as in Lemma 4.62. If $\hat{\phi}/\!\!/\hat{\sigma} = \hat{\phi}$, then $\bar{\phi}$ does not eliminate $\hat{\sigma}$.*

Proof As for Lemma 4.29, noting that the proof of Proposition 4.27 on which it is based, only needs to be supplemented by the observation that canonisation does not change the lengths of outermost developments as established in Lemma 4.64. □

[13] In the definition of a development either the orthogonal or the weakly orthogonal projection may be employed.

Finally, the Head-Cluster and Non-Head-Cluster Lemmas allow to prove the main result of this section, by adapting the alternative proof to Theorem 3.17 as presented on page 31.

Theorem 4.66 *Weakly head normalising weakly orthogonal TRSs are acyclic.*

Proof For a proof by contradiction, assume there would exist a non-empty parallel cycle. Select among the non-empty parallel cycles of minimal length a parallel head cycle, say σ starting with head step ψ from t. For any multi-step ϕ co-initial to σ, Lemma 4.62 yields an extension $\hat{\phi}$ of ϕ and a co-initial parallel head-cluster cycle $\hat{\sigma}$ of the same length as σ starting with a head-cluster step $\hat{\psi}$. By the Head-Cluster Lemma $\hat{\psi}/\!\!/\hat{\phi}$ is either empty or a head-cluster step again. But the former cannot hold, since then $\hat{\sigma}/\!\!/\hat{\phi}$ would be a parallel cycle shorter than σ, which by the Non-Head-Cluster lemma is non-empty, contradicting minimality of the length of σ. Therefore, $\hat{\sigma}/\!\!/\hat{\phi}$ is a parallel head-cluster cycle again. Now by the weak head normalisingness assumption there is a reduction τ from t to head normal form s. Hence, by Lemma 2.9 and the above, we may assume that $\tau/\sigma = \tau$ and that σ/τ is a parallel head-cluster cycle. But this would contradict s being in head normal form. $\square$

Remark 4.67 In [20] it was shown that a term having a head *loop* is not normalising. Theorem 4.66 generalises this from loops to cycles and from normalisation to head normalisation, but requires the latter for *all* terms.

5 Acyclicity of weakly head normalising λ-calculi

In this section, we show that reductions are acyclic in λ-calculi which are weakly head normalising. In particular, we show this for the ordinary $\lambda\beta$-calculus, the extensional $\lambda\beta\eta$-calculus, and two $\lambda\beta$-calculi with explicit substitutions. The proof of the former is an easy corollary to the earlier results.

Theorem 5.1 *For sub-calculi of $\lambda\beta$-calculus W(H)N $\Rightarrow$ AC.*

Proof This is just Examples 3.6 and 4.4. $\square$

Next, we turn our attention to the $\lambda\beta\eta$-calculus.

Theorem 5.2 *Head strategies are hyper-normalising for $\lambda\beta\eta$-calculus.*

Proof Let t be a term which has a normal form. By η-postponement [4], we may assume the reduction σ to normal form is of shape $\sigma_\beta \cdot \sigma_\eta$, for some β-reduction $\sigma_\beta{:}t \twoheadrightarrow s$ and η-reduction $\sigma_\eta{:}s \twoheadrightarrow u$. Moreover, we may assume that s is in β-normal form (if not, performing a maximal β-reduction from s terminates and leaves a residual η-reduction to u, as one easily checks).

We will show that projecting such a reduction over a hyper-head reduction τ eventually results in an empty reduction. To that end, we measure σ by first the number of η-redexes contracted in σ_η, and second the μ-measure [18] of σ_β. We distinguish cases on whether the (first) step $\phi{:}t \to t'$ of τ projected over, using the weakly orthogonal projection of [18,17], has a residual after σ_β or not.

In case it does, ϕ must be an η-step since σ_β was assumed to end in β-normal form. Moreover, ϕ must be eliminated by σ_η since the latter was assumed to end in η-normal form. Thus the projection of σ_η contracts one η-redex less.

In case it does not, ϕ must have been eliminated by σ_β. But then the μ-measure of the projection of σ_β has been decreased (in case of overlap) or is unchanged (in case all residuals of ϕ have been erased).

Noting that a head step can not be erased we conclude. □

Theorem 5.3 *For sub-calculi of the $\lambda\beta\eta$-calculus, WN $\Rightarrow$ AC.*

Proof By Theorem 5.2. □

We do not know whether the assumption of the theorem can be weakened from WN to WHN. The following example displays two problems which show up in a proof attempt along the lines above.

Example 5.4 The leftmost outermost strategy need not be head normalising, as witnessed by the term $t = \lambda x.z(ss)x$ with $s = \lambda y.Kyxy$. The leftmost outermost strategy reduces this term to itself, whereas the head normal form $z\Omega$ can be reached by first contracting the K-redexes.

A head normal form can sometimes be reached without performing a head step, as witnessed by the term $t = \lambda x.Ix$ which reduces by one non-head-β-step to the (head) normal form $\lambda x.x$ (it could also be reduced to that head normal form by a head-η-step, but that's not the issue).

Finally, we turn our attention to the two $\lambda\beta$-calculi with explicit substitutions $\lambda\mathtt{x}^-$ and $\lambda\sigma$. The proof of acyclicity of weakly head normalising such is by reduction to acyclicity of the associated weakly head normalising $\lambda\beta$-calculus.

First, we consider the $\lambda\mathtt{x}^-$-calculus [7], which is the $\lambda\beta$-calculus extended with an explicit substitution operator $[:=]$ with rules, for x distinct from y:

$$\begin{aligned}
(\lambda x.M)N &\to M[x{:=}N]\\
(M_1M_2)[x{:=}N] &\to M_1[x{:=}N]M_2[x{:=}N]\\
(\lambda x.M)[y{:=}N] &\to \lambda x.M[y{:=}N]\\
x[x := N] &\to N\\
y[x := N] &\to y
\end{aligned}$$

Theorem 5.5 *Head reductions are hyper head normalising in $\lambda\mathtt{x}^-$.*

Proof We have to show that reductions that eventually always perform a head step, are head normalising. Hence, for a proof by contradiction, assume that t allows both for a reduction σ containing infinitely many head steps as well as a reduction τ to head normal form s. We claim that we may assume that σ contains infinitely many Beta-steps which project onto head-β-steps. The result follows, since then in the $\lambda\beta$-calculus, the projection $\sigma{\downarrow_{\mathtt{x}}}$ is a reduction containing infinitely many head-β-steps, and the projection $\tau{\downarrow_{\mathtt{x}}}$ is a reduction to the head normal form s, since $\lambda\mathtt{x}^-$-head normal forms are $\lambda\beta$-normal forms. This combination is impossible due to Theorem 3.4.

We prove the claim in four proof steps.

First, since $\lambda\mathtt{x}^-$ is a left-linear second-order PRS the standardisation theorem holds for it [18]. That is, any reduction can be transformed into a reduction which is standard in the sense that a redex-pattern contracted in any of its steps overlaps the redex-pattern of the first step (if any) after and outside or to the left of it. Since $\lambda\mathtt{x}^-$ is left-normal, the standard reduction can be obtained by repeatedly permuting so-called anti-standard pairs, i.e. two consecutive steps such that the redex-pattern contracted by the latter is entirely above or to the left of the former [12]. Hence, standardising σ yields a standard reduction still containing infinitely many head steps. Therefore, we may assume without loss of generality that σ is standard.

Second, all Beta-steps of σ must be on the spine. This holds true, since by standard-ness the position of a contracted redex-pattern ϕ must overlap the redex-pattern of the first redex-pattern ψ contracted outside it in σ. Such a step ψ always exists since σ contains infinitely many head steps. From the form of the left-hand sides of the $\lambda\mathtt{x}^-$-rules, we have that if ψ is on the spine, then ϕ must be so as well. Since head steps are on the spine, we conclude that all of them are.

Third, there must always eventually be a Beta-step in σ. This holds by SN of the explicit substitution rules of $\lambda\mathtt{x}^-$, and infiniteness of σ.

Finally, assuming a Beta-step ϕ on the spine in σ, consider a non-substitution symbol above it (not necessarily properly so), hence on the spine, which is closest to the head. Note that by the form of the rewrite rules the symbol has a unique residual, until contracted as part of a Beta-redex, which is then closest to the head and on the spine again. Moreover, after a *head* step the residual is closer to the head. Hence, eventually it must be contracted. $\square$

Remark 5.6 One could prove that in a $\lambda\mathtt{x}^-$-reduction containing infinitely many head steps, infinitely many of these must in fact be Beta-steps. We refrain from doing so here, since it does not generalise to the $\lambda\sigma$-calculus.

From hyper head normalisation we have acyclicity of $\lambda\mathtt{x}^-$, as before.

Theorem 5.7 *If $\lambda\mathtt{x}^-$ is weakly head normalising, then it is acyclic.*

As before, this result in itself is not interesting as the $\lambda\mathtt{x}^-$-calculus is not weakly head normalising, but it does become interesting for sub-calculi. Let $\mathcal{X}$ be a sub-calculus of $\lambda\mathtt{x}^-$ in the sense that it is closed under reductions and taking sub-terms. In order to be able to reduce acyclicity of $\mathcal{X}$ to that of a

$\lambda\beta$-calculus, it suffices to define the $\lambda\beta$-calculus induced by $\mathcal{X}$, and verify that weak (head) normalisation is preserved.

Definition 5.8 The sub-ARS of the $\lambda\beta$-calculus *induced by* $\mathcal{X}$ is $\to_\beta$ restricted to terms in $\mathcal{X}$.

Proposition 5.9 *For the sub-ARS of* $\to_\beta$ *induced by* $\mathcal{X}$ *it holds:*

- *it is a sub-calculus of* $\to_\beta$*, and*
- *it is weakly (head) normalising, if* $\to_{\lambda\mathtt{x}^-}$ *is so on* $\mathcal{X}$*.*

Proof

- Note that any β-step on λ-terms may be simulated by a number of $\lambda\mathtt{x}^-$-steps [11, Lem. 5(3)]. Thus, closure under taking sub-terms and under β-reduction, follow from the corresponding properties of $\mathcal{X}$.
- Let t be a λ-term. By assumption, it can be reduced to some (head) normal t'. Define $s = t'{\downarrow_{\mathtt{x}}}$, that is, s is the substitution normal form [11, Lem. 5(1)] of t'. By [11, Lem. 5(1,2)], s is a λ-term which is reachable by β-steps from t. Refining the observation of the first item to: a (head) β-step from a λ-term gives rise to a (head) Beta-step from it, we conclude that t is β-reducible to a (head) normal form, since s being a $\lambda\mathtt{x}^-$-reduct of a (head) normal form, s itself is one too.

$\square$

Hence, Theorem 5.7 extends to sub-calculi of $\lambda\mathtt{x}^-$. Next, we consider the calculus of explicit substitutions $\lambda\sigma$ as introduced in [1] (see e.g. [17, Sec. 3.6]).

Theorem 5.10 *Head strategies are hyper head normalising for* $\lambda\sigma$*.*

Proof This follows from [15, Thm. 3].

Here we present a direct proof, by adapting the earlier proof for $\lambda\mathtt{x}^-$ above, only indicating the essential properties used in its four proof steps.

First, standardisation holds for $\lambda\sigma$ as it is a left-linear left-normal TRS (see Figure 19), so we may assume reductions containing infinitely many head steps to be standard.

Second, all the Beta-steps must be on the spine. This follows by proving the stronger property that all sub-terms above a Beta-redex-pattern (in particular the term itself), must in fact have type *Expression*, which holds by standardness since a Beta-redex has type *Expression*, terms of type *Expression* only overlap terms having that type again, and if a sub-term on the spine is overlapped by a sub-term of type *Expression*, then the latter is on the spine again (see Figure 19).

Third, the explicit substitution calculus of $\lambda\sigma$ is strongly normalising.

Checking the fourth step proceeds as for $\lambda\mathtt{x}^-$, verifying that an outermost application symbol on the spine is preserved by all steps until contracted by a Beta-redex, and we conclude. $\square$

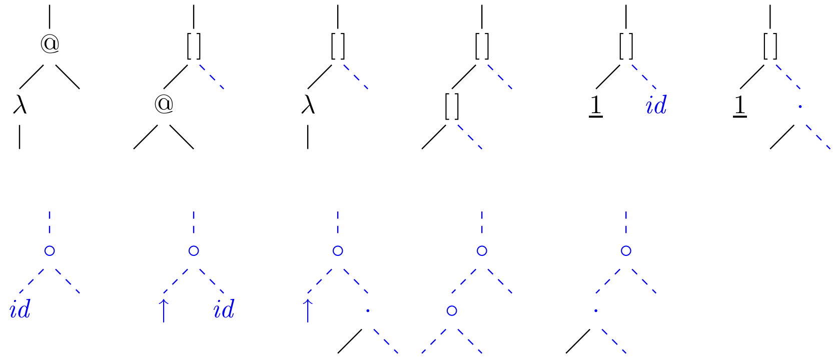

Figure 19. Left-hand sides of $\lambda\sigma$-rules, with *Substitutions* dashed/in blue

As an immediate corollary we have.

Theorem 5.11 *Weakly (head) normalising sub-calculi of $\lambda\sigma$ are acyclic.*

One way of construing these results is as an explanation for the fact that the counter-examples to preservation of strong normalisation for typed λ-calculi with explicit substitution, as found in the literature starting from the seminal paper [14], display acyclic behaviour and hence unbounded growth; it *couldn't* have been otherwise! More precisely, we have:

Corollary 5.12 *Typed $\lambda\sigma$ is acyclic.*

Proof Let t be a typed $\lambda\sigma$-term and consider its family, which is a sub-calculus (see Example 3.6). Any term s in this sub-calculus is seen to be weakly head normalising as follows.

First reduce s to substitution normal form. This yields a term consisting of a (possibly empty) context of sub-terms of type *Substitution* having arguments of type *Expression* which are (translations of) typed $\lambda\beta$-terms. Since typed $\lambda\beta$-calculus is SN, these arguments can be reduced to normal form in the $\lambda\sigma$-calculus, by simulating the reduction to normal form in the $\lambda\beta$-calculus. This yields a term in $\lambda\sigma$-normal form.

Therefore, by the previous theorem, t is acyclic. □

It is not immediately clear how to generalise this to the untyped case. In particular, starting from some untyped $\lambda\beta$-term t which is strongly normalising, is its associated $\lambda\sigma$-term acyclic? The point is that it is not clear why its family should be weakly head normalising: following the procedure in the proof, the arguments in substitution normal form might be $\lambda\beta$-terms which are *not* in the $\lambda\beta$-family of t!

Remark 5.13 It is easy to prove that the $\lambda\beta$-family of a $\lambda\beta$-term is the same as the restriction of its $\lambda\mathtt{x}^-$-family to $\lambda\beta$-terms; prove by induction on the length of the $\lambda\mathtt{x}^-$-reduction that the substitution normal form of any term in a closure is reachable. For λ-calculi having some form of composition, like $\lambda\sigma$, this proof method does not work immediately, because of the problem mentioned above. Still, we conjecture that also for these, the property holds. Since the $\lambda\sigma$-calculus is rather unwieldy, having 11 rules, it might be simpler to attempt to prove the conjecture first for the $\lambda\mathtt{x}^-$-calculus extended with the single *substitution lemma* rule

$$M[x{:=}N][y{:=}P] \to M[y{:=}P][x{:=}N[y{:=}P]]$$

The problem in a proof attempt is then caused by the sub-term $N[y{:=}P]$ which might not occur in the substitution normal of the term, i.e. not in the corresponding $\lambda\beta$-term, in case x is erased.

Intuitively, the result should hold because of standardisation: the sub-term can only be problematic if the left-hand side is caused by inside-out creation. But then the application of the substitution lemma rule, transforms it into an outside-in creation and there should be an outside-in reduction creating the sub-term as well. It remains to formalise this intuition.

Vice versa, the λ-calculi with explicit substitutions can be construed as counter-examples of sorts to the Barendregt–Geuvers–Klop Conjecture 1.1. As their failure to be strongly normalising, in the typed case, can be attributed to their ability to express and manipulate 'composition of substitution', it would be interesting to formulate this ability abstractly and prove that ordinary typed $\lambda\beta$-calculi are lacking it.

6 Conclusions

We have generalised our earlier result, WN $\Rightarrow$ AC for orthogonal TRSs, into several directions. The generalisations to rewriting systems which still have a nice theory of permutation, were relatively straightforward, consisting in generalising the Head Normalisation Theorem to these cases. In contrast, the generalisations to rewriting systems which do not have a nice theory of permutation, in particular to weakly orthogonal and to non-fully-extended rewriting systems, and to the λ-calculi with explicit substitutions $\lambda\mathtt{x}^-$ and $\lambda\sigma$, required substantial effort. The reason is that generalising the Head Normalisation Theorem is not straightforward for λ-calculi with explicit substitutions, and even fails for weakly orthogonal and non-fully extended rewriting systems. To overcome these problems, we have introduced several new notions and techniques, the main ones of which were:

- the full-extension of a non-fully-extended PRS (Section 3.1),
- a theory of projections of cycles, in analogy to the theory of repeating decimals (Section 4.1), and
- a theory of clusters of redex-patterns in weakly orthogonal term rewriting systems, showing that they are coverable by chains (pages 33–40).

This paper is but a small step towards a theory of cyclic reductions. The latter and, more generally, the theory of non-terminating reductions remain largely unexplored. Apart from the Barendregt–Geuvers–Klop Conjecture 1.1, we have left several concrete problems open. We conclude by presenting these as the following conjectures:

- WHN $\Rightarrow$ AC, for weakly orthogonal PRSs (Conjecture 1.2).
- Weakly orthogonal TRSs do not constitute a residual system (Remark 2.26).
- Head strategies are hyper-head-normalising for the $\lambda\beta\eta$-calculus (Example 5.4).
- The image in an explicit substitution calculus $\lambda?$ of the $\lambda\beta$-family of a $\lambda\beta$-term, is the same as the restriction of its $\lambda?$-family to images of $\lambda\beta$-terms (Remark 5.13).

Acknowledgments

The third author would like to thank the members of the Computing Science Department (InCo) of the Universidad de la Republica, Montevideo, Uruguay, and Alejandro Rios for their hospitality while on sabbatical leave.

References

[1] Abadi, M., L. Cardelli, P.-L. Curien and J.-J. Lévy, *Explicit substitutions*, Journal of Functional Programming **1** (1991), pp. 375–416.

[2] Allen, J., *Maintaining knowledge about temporal intervals*, Communications of the ACM **26** (1983), pp. 832–843.

[3] Baader, F. and T. Nipkow, "Term Rewriting and All That," Cambridge University Press, 1998, xii+301 pp.

[4] Barendregt, H., "The Lambda Calculus: Its Syntax and Semantics," Studies in Logic and the Foundations of Mathematics **103**, Elsevier, 1985, second edition.

[5] Barendregt, H., J. Bergstra, J. Klop and H. Volken, *Degrees, reductions and representability in the lambda calculus* (1976), the Blue Preprint.

[6] Bethke, I., J. Klop and R. de Vrijer, *Descendants and origins in term rewriting*, Information and Computation **159** (2000), pp. 59–124.

[7] Bloo, R., "Preservation of Termination for Explicit Substitution," Ph.D. thesis, Technische Universiteit Eindhoven (1997).

[8] de Vrijer, R., "Surjective Pairing and Strong Normalization: Two Themes in Lambda Calculus," Ph.D. thesis, Universiteit van Amsterdam (1987).

[9] Huet, G. and J.-J. Lévy, *Computations in orthogonal rewriting systems, Part I + II*, in: J. Lassez and G. Plotkin, editors, *Computational Logic – Essays in Honor of Alan Robinson* (1991), pp. 395–443, revised version of: Call-by-need computations in non-ambiguous linear term rewriting systems, Rapport 359, INRIA, 1979.

[10] Ketema, J., J. Klop and V. v. Oostrom, *Vicious circles in orthogonal term rewriting systems*, in: S. Antoy and Y. Toyama, editors, *Proceedings of the workshop on reduction strategies, WRS 2004*, Electronic Notes in Theoretical Computer Science (200x), to appear.

[11] Khasidashvili, Z., M. Ogawa and V. van Oostrom, *Uniform normalisation beyond orthogonality*, in: A. Middeldorp, editor, *Proceedings of the Twelfth International Conference on Rewriting Techniques and Applications (RTA '01)*, Lecture Notes in Computer Science (2001), pp. 122–136.

[12] Klop, J., "Combinatory Reduction Systems," Mathematical Centre Tracts **127**, Mathematisch Centrum, Amsterdam, 1980.

[13] Mayr, R. and T. Nipkow, *Higher-order rewrite systems and their confluence*, Theoretical Computer Science **192** (1998), pp. 3–29.

[14] Melliès, P.-A., *Typed lambda-calculi with explicit substitutions may not terminate*, in: M. Dezani-Ciancaglini and G. Plotkin, editors, *Proceedings of the Second International Conference on Typed Lambda Calculi and Applications (TLCA '95)*, Lecture Notes in Computer Science **902** (1995), pp. 328–334.

[15] Melliès, P.-A., *Axiomatic rewriting theory II, the $\lambda\sigma$-calculus enjoys finite normalisation cones*, Journal of Logic and Computation **10** (2000), pp. 461–487.

[16] O'Donnell, M., "Computing in Systems Described by Equations," Lecture Notes in Computer Science **58**, Springer, 1977.

[17] Terese, "Term Rewriting Systems," Cambridge Tracts in Theoretical Computer Science **55**, Cambridge University Press, 2003.

[18] van Oostrom, V., *Normalisation in weakly orthogonal rewriting*, in: P. Narendran and M. Rusinowitch, editors, *Proceedings of the Tenth International Conference on Rewriting Techniques and Applications (RTA '99)*, Lecture Notes in Computer Science **1631** (1999), pp. 60–74.

[19] van Oostrom, V., *Delimiting diagrams* (2004), University of Utrecht, Draft, `http://www.phil.uu.nl/~oostrom/publication/pdf/dd.pdf`.

[20] van Oostrom, V., *Trivial* (August 1998), University of Utrecht, Draft, `http://www.phil.uu.nl/~oostrom/publication/pdf/trivial.pdf`.