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 [7], 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 [7]) 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 [7], 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 [7] 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 [3]. 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. The combination of all three extensions remains open however:

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

Finally, we turn our attention to λ-calculi, in particular to $\lambda\beta$-calculi, $\lambda\beta\eta$-calculi, yielding the same partial answer to the Barendregt–Geuvers–Klop conjecture for typed $\lambda\beta\eta$-calculi as for such calculi without η, and the λ-calculi with explicit substitutions $\lambda\mathtt{x}^-$ and $\lambda\sigma$.

This is an extended abstract of the technical report [6]. Because of space limitations we can present only two of the four novel proof techniques introduced there. We have chosen to completely develop here the proof techniques of *fully extending a rewriting system* and *projection of cycles along extensions*, referring the reader to the technical report for details of the other proof techniques (*covering clusters of overlapping redex-patterns by chains* and *eliminating redex-patterns from clusters*). For the same reason, we assume here familiarity with the basic theory of confluence for (weakly) orthogonal TRSs [14, Chapter 4].

Throughout, the rewriting systems considered are left-linear.

2 Acyclicity for separate generalisations

We generalise the WN $\Rightarrow$ AC result and its proof, as obtained in [7] for orthogonal first-order term rewriting systems, into each of the three individual directions mentioned in the introduction. To that end, we first recapitulate their main ingredients as presented in [7].

See [14,2] for introductions to abstract rewriting systems (ARSs). Concretely, we use the ARSs of [14, Definition 8.2.2], a, b, c, ... to denote objects, ϕ, ψ, χ, ... to denote steps, and σ, τ, υ, ... to range over reductions.

Definition 2.1 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. For example, the TRS in the introduction is AC.

Definition 2.2 A step $C[l^\sigma] \to C[r^\sigma]$ according to rule $l \to r$ is a *head* step if C is the empty context. A set of reductions is said to *have sub-reductions*, if for any non-head reduction in the set with source t, there is a reduction in the set having some proper subterm of t as source.

Notions for ordinary reduction relativise in the obvious way to head reduction, e.g. a TRS is WHN (weakly head normalising) if for every term there is a reduction to a term in head normal form.

Theorem 2.3 (Head normalisation) *If a term allows a reduction containing infinitely many head steps in an orthogonal TRS, a so-called* hyper-head *reduction [14, Definition 9.1.18], then it does not have a normal form.*

The (contrapositive of) the statement of the theorem may be abbreviated to *head reductions are hyper-normalising* or equivalently *hyper-head reductions are normalising.*

Lemma 2.4 (Sub-cycle) *Cycles have sub-reductions in TRSs.*

Theorem 2.5 ([7]) *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 Sub-Cycle Lemma and 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 normal form, contradicting the assumption that $\mathcal{R}$ was weakly normalising.□

The proof is abstract in that it only depends on the Head Normalisation Theorem and the Sub-Cycle Lemma. Hence it will suffice to generalise these.

For the generalisation of the implication WN $\Rightarrow$ AC to WHN $\Rightarrow$ AC, only the Head Normalisation Theorem needs to be generalised.

Theorem 2.6 *Weakly head normalising orthogonal TRSs are acyclic.*

Proof To reach a contradiction with the weak head normalisation assumption, as before, it suffices to prove that head reductions are even hyper-*head* normalising for orthogonal TRSs. Observing that head steps are head needed in the sense of [14, Definition 9.5.52], i.e. they cannot be eliminated by other non-overlapping steps, the result follows from the fact that head needed reductions are hyper-head normalising for orthogonal TRSs ([14, Theorem 9.2.60]). □

Example 2.7 Replacing the right-hand side of the rule $f(x) \to b$ by $c(a)$ transforms the TRSs of the introduction into an orthogonal TRS which is WHN (but no longer WN). We conclude AC from the theorem.

To generalise the implication WN $\Rightarrow$ AC from orthogonal to weakly orthogonal TRSs, that is to left-linear TRSs the critical pairs of which are trivial (c.f. Example 2.9), it suffices to adapt the Head Normalisation Theorem.

Theorem 2.8 *Weakly normalising weakly orthogonal TRSs are acyclic.*

Proof To reach a contradiction, reasoning as before, it suffices to prove that head reductions are hyper-normalising also for *weakly* orthogonal TRSs. Observing that head steps are outermost steps, and hence that a reduction which eventually always performs a head step is outermost fair in the sense of [14, Definition 9.3.1], i.e. outermost redexes are eventually eliminated, the result follows from the fact that outermost-fair reductions are normalising for weakly orthogonal TRSs ([12, Theorem 1], [14, Theorem 9.3.12]). □

Example 2.9 The predecessor/successor TRS with rules $P(S(x)) \to x$ and $S(P(x)) \to x$ is weakly orthogonal; both its critical pairs are trivial:

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

We conclude AC from WN (it is even SN, but that's not the point here).

Our third generalisation of the implication WN ⇒ AC, from first-order to higher-order term rewriting systems (PRSs [14, Chapter 11]), is more difficult to establish. We will first prove it for PRSs which are *fully extended* or *fully applied*, 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.

Example 2.10 The paradigmatic example of a non-fully-applied PRS is the $\lambda\beta\eta$-calculus, with rules, for M, N terms:

$$\beta{:}(\lambda x.M)N \to M[x{:=}N] \qquad\qquad \eta{:}\lambda x.Mx \to M, \text{ if } x \notin M$$

Whereas the β-rule is fully applied, the η-rule is not, as its applicability depends on whether the bound variable occurs. In general, a rule scheme containing variable conditions on its terms yields a non-fully-applied PRS rule.

The problem with non-fully-applied rules is that the proof techniques employed above, normalisation of head needed/outermost-fair strategies, fail.

Example 2.11 Consider the term $t = f(x.g(e(x)))$ in the orthogonal PRS with rules $f(x.Z) \to a$, $g(Z) \to g(Z)$, and $e(Z) \to a$, where Z is a meta-variable of base type. The f-rule is non-fully-applied as it tests for the absence of the bound variable x in its argument Z (in PRSs, instantiating the meta-variable Z in the left-hand side $f(x.Z)$ of the first rule requires that no variable be captured, so no x may occur in the instance of Z). Since this test fails, t is not yet a redex, and hence the g-redex is the outermost redex in t. On the one hand, repeatedly contracting this outermost redex gets you nowhere. On the other hand, a reduction to normal form is possible, by first contracting the *non-head-needed* and *non-outermost* e-redex in t which *erases* the variable x, yielding the term $f(x.g(a))$ which reduces in one step to the normal form a.

Note that the example is not a counter-example to the implication WN ⇒ AC, since the sub-term $g(e(x))$ of t does not have a normal form.

Theorem 2.12 *A weakly normalising fully applied orthogonal PRS is acyclic.*

Proof We generalise the Head Normalisation Theorem and the Sub-Cycle Lemma. For the former, we make an appeal to [12, Theorem 1], expressing normalisation of outermost-fair strategies for to fully applied orthogonal PRSs.

For the latter this is trivial. In fact, cyclic reductions are closed under sub-reductions for arbitrary PRSs. □

Example 2.13 Since the $\lambda\beta$-calculus (without the η-rule) is fully applied and orthogonal, we conclude that if it were normalising, then it would be acyclic. This implication is trivial since the $\lambda\beta$-calculus obviously is *not* weakly normalising. Still, the proof is readily seen to go through for any sub-calculus of a fully applied orthogonal PRS which is closed under reduction. In fact, the $\lambda\beta$-calculus itself is only a sub-calculus of its PRS encoding, as the latter allows too many terms. However, the examples of sub-calculi important to us are typed $\lambda\beta$-calculi. Weakly normalising such are AC by the theorem.

To generalise Theorem 2.12 to non-fully-applied PRSs, note that the problem in Example 2.11 is that the reduction step erasing the variable is *intuitively* needed to create the head redex-pattern, but it is not needed in the technical sense, since the step does not 'contribute' to the redex-pattern via the ordinary descendant relation. We construct the full-extension $\bar{\mathcal{H}}$ of a PRS $\mathcal{H}$ where this problem is resolved, while preserving the good properties of $\mathcal{H}$,

Definition 2.14 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.[4] 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.

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 of an $\bar{l}$-redex-pattern serve to erase variables restricted by variable conditions in $\mathcal{H}$.

Example 2.15 Fully-extending the $\lambda\beta\eta$-calculus does not change the β-rule. However, the η-rule changes into (the PRS representation of):

$$\bar{\eta}{:}\lambda x.Mx \to M[x{:=}a]$$

for an arbitrary closed term a. Note the absence of the usual side-condition on M. For instance, 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).

Lemma 2.16 *Let $\bar{\mathcal{H}}$ be the full-extension of the PRS $\mathcal{H}$. Then*

- $\to_{\mathcal{H}} \subseteq \to_{\bar{\mathcal{H}}}$,

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

- *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 See [6]. □

Theorem 2.17 *Weakly normalising orthogonal PRSs are acyclic.*

Proof [Sketch] It suffices to generalise the Head Normalisation Theorem. So consider a reduction σ in $\mathcal{H}$ which eventually always performs a head step. By Lemma 2.16, σ is a reduction in $\bar{\mathcal{H}}$ as well, having the same property. We claim that again [12, Theorem 1] can be applied, yielding that σ is a normalising $\bar{\mathcal{H}}$-reduction. Therefore, by Lemma 2.16 again, the reduction will find a $\mathcal{H}$-head normal form as well.

To prove the claim it suffices that the proof of [12, Theorem 1] generalises from fully applied orthogonal PRSs to $\bar{\mathcal{H}}$. Apart from orthogonality, its proof requires the following properties, all of which are implied by Lemma 2.16:[5]

- residuals of redex-patterns are redex-patterns again,
- head normal forms are closed under reduction,
- reductions are closed under standardisation, and
- head normal forms are closed under expansions below the head.

□

Example 2.18 The $\lambda\eta$-calculus (without β) is an orthogonal PRS, hence we may conclude its acyclicity from its weak normalisation.

3 Acyclicity for the pair-wise combinations

We consider the three pair-wise combinations of the three generalisations in the previous section of the implication WN $\Rightarrow$ AC. Two of them are easy.

Theorem 3.1 *For weakly orthogonal fully applied PRSs, WN $\Rightarrow$ AC.*

Proof As for Theorem 2.8: [12, Theorem 1] applies to fully applied PRSs. □

Remark 3.2 We don't know whether fully-applied-ness can be dropped here.

Theorem 3.3 *For orthogonal PRSs, WHN $\Rightarrow$ AC.*

Proof As for Theorem 2.17: its proof established that the set of $\bar{\mathcal{H}}$-head normal forms is a set of *results*, as required for head normalisation [12, Theorem 2], instead of just normalisation [12, Theorem 1]. □

Example 3.4 WHN sub-calculi of the $\lambda\beta$-calculus are AC.

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

The third combination seems to evade standard theory.

Theorem 3.5 *For weakly orthogonal TRSs, WHN $\Rightarrow$ AC.*

Proof See [6] and the next section. $\square$

4 Acyclicity by projection of cycles

We present an alternative proof of Theorem 2.6 based on the projection of cycles. Apart from that we think the concepts involved are interesting in their own right, the reason for presenting this alternative proof technique here is that it is at the basis of our proof of Theorem 3.5, as presented in [6], a theorem which resisted all other methods. As before we present the proof method in an abstract fashion, relying instead of on the Head Normalisation Theorem, on the Head, the Non-Head, and the Extension Lemma, respectively.

4.1 Head

We prove the Head Lemma, showing that head steps project in a particularly simple way. Here we say that an ARS has projection if it has the diamond property [14, Definition 1.1.8(v)], and if moreover that fact is witnessed by a projection function (see Figure 2).

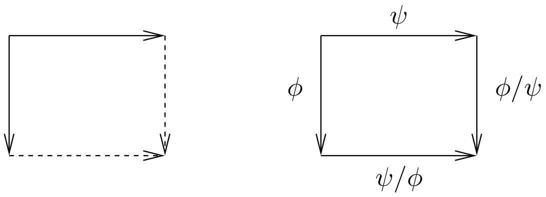

Figure 2. Diamond property and projection

Definition 4.1 An ARS $\to$ is said to *have projection* if there is a function $/$ from pairs of co-initial steps ϕ, ψ to steps such that:

- 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*.

Any ARS having the diamond property can be turned into an ARS having projection by *choosing* for every pair of co-initial steps a particular 'diamond completing it'. It is a basic fact [14, Chapter 4] that for an orthogonal TRS $\mathcal{R}$ multi-steps $\mathrel{\circ\!\!\!\!\longrightarrow}_{\mathcal{R}}$ and parallel steps $\mathrel{+\!\!\!\!\!\longrightarrow}_{\mathcal{R}}$ [14, Figure 8.8], allowing the simultaneous contraction of an arbitrary number of redex-patterns (at disjoint positions), have the diamond property. In fact, a projection function $/$ working both for $\mathrel{\circ\!\!\!\!\longrightarrow}_{\mathcal{R}}$ and for $\mathrel{+\!\!\!\!\!\longrightarrow}_{\mathcal{R}}$ was presented in [14, Definition 8.7.4].

Lemma 4.2 *Residuals of parallel steps after multi-steps are parallel steps.*

Proof Using that $\mathrel{\Vert\!\!\!\!\!\!\longrightarrow}_{\mathcal{R}}$ has projection, it suffices to note that any multi-step can be developed into a series of parallel steps. □

A multi-step (parallel step) is empty, if it contracts an empty set of redexes.

Lemma 4.3 (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.*

Proof Trivial. □

4.2 Non-Head

We prove the Non-Head Lemma, showing that if a head parallel cycle leaves no (an empty) residual after a multi-step, then the multi-step cannot have itself as residual after the head parallel cycle. Here a *head parallel cycle* is a cycle of $\mathrel{\Vert\!\!\!\!\!\!\longrightarrow}$-steps the first of which is a head step. We first lift projection from steps to reductions and cycles by means of *tiling* (see Figure 3).

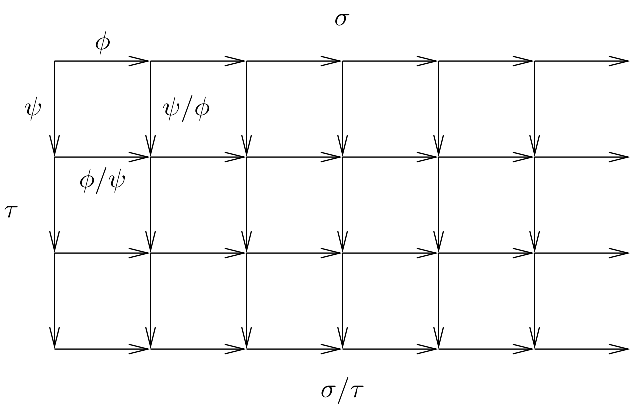

Figure 3. Projection of infinite reduction σ over finite reduction τ by tiling

Definition 4.4 Let $\to$ be an ARS having projection $/$. Then the projection of a finite or infinite reduction over a finite one is defined by (cf. [5]):

$$\sigma/\emptyset = \sigma \qquad\qquad \emptyset/\psi = \emptyset$$

$$\sigma/(\psi\cdot\tau) = (\sigma/\psi)/\tau \qquad\qquad (\phi\cdot\sigma)/\psi = (\phi/\psi)\cdot(\sigma/(\psi/\phi))$$

Thus, $\twoheadrightarrow$ has projection if $\to$ has, corresponding to the well-known fact that $\to$ is confluent ($\twoheadrightarrow$ has the diamond property) if $\to$ has the diamond property. Note that the residual of σ always has the same length as σ. To prove the Non-Head-Lemma, it suffices to define a measure on multi-steps which decreases along the projection of the latter over head parallel cycles.

Definition 4.5 The measure $\mu(\phi)$ of a multi-step ϕ is the maximal length of an outermost development of ϕ, where the *outermost development* of ϕ is:

- 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 [14, 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 in the sense of [13].

Example 4.6 Consider in the TRS with rules $a \to b$ and $f(x,y) \to g(x,x)$, 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 $t = f(f(a,a),f(a,a))$. An outermost development of t^* is:

$$\underline{f}(f(a,a),f(a,a)) \to g(\underline{f}(a,a),f(a,a)) \to g(g(\underline{a},a),f(a,a)) \to$$
$$g(g(b,\underline{a}),f(a,a)) \to g(g(b,b),\underline{f}(a,a)) \to g(g(b,b),g(\underline{a},a)) \to$$
$$g(g(b,b),g(b,\underline{a})) \to g(g(b,b),g(b,b))$$

Hence $\mu(\phi) = 7$. (Note that there is an *ordinary* development of length 13).

If ψ/ϕ is empty, then we say that ψ *eliminates* or *absorbs* ϕ.

Proposition 4.7 *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 One shows that for any multi-step ϕ with ψ/ϕ empty and any outermost development of ϕ/ψ, one may construct an outermost development of ϕ, of at least the same length, by induction on $\mu(\phi)$. $\square$

Remark 4.8 To see that Proposition 4.7 fails if we were to take the maximal length of an *ordinary* development as measure of a multi-step, consider the multi-step ϕ:$\underline{e}(d(\underline{a})) \multimap c$ and the co-initial step ψ:$e(\underline{d}(a)) \multimap e(f(a,a))$ in the orthogonal TRS with rules $a \to b$, $e(x) \to c$, and $d(x) \to f(x,x)$.

Lemma 4.9 (Non-Head) *If in an orthogonal TRS, a head parallel cycle σ is co-initial to a multi-step ϕ with $\phi/\sigma = \phi$, then ϕ does not eliminate σ, that is, the cycle σ/ϕ does not consist solely of empty steps.*

Proof Suppose (the successive residuals of) ϕ would eliminate all parallel steps of σ. Then, by repeated application of Proposition 4.7, $\mu(\phi) > \mu(\phi/\sigma) = \mu(\phi)$, since the first step of σ was assumed to be a head step. $\square$

4.3 Extension

Recall that the goal was to give a proof of WHN $\Rightarrow$ AC, based on the projection of cycles. However, the following example shows that the projection of a cyclic reduction over a multi-step needs *not* yield a cyclic reduction again.

Example 4.10 The step σ:$\underline{f}(a,a,a) \to f(a,a,a)$ is a cycle in the TRS with rules $a \to b$ and $f(x,y,z) \to f(y,x,a)$. However, the residual of σ after the parallel step ϕ:$f(\underline{a},a,\underline{a}) \twoheadrightarrow\!\!\!\!\!\!+\; f(b,a,b)$ is $\underline{f}(b,a,b) \to f(a,b,a)$ which obviously is *not* a cycle (see Figure 4 left).

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

Figure 4. Residuals of the cycle σ along the step ϕ and its extension $\bar{\phi}$, respectively.

Yet, there always is an *extension* of the multi-step which preserves cyclicity.

Definition 4.11 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 *extends* ϕ. An *extension* (function) maps each step to some extension of it.

Intuitively, one may think of extending as one reduction *overtaking* the other. Obviously, it is reflexive and transitive, i.e. a quasi-order.

Example 4.12 The parallel step $\bar{\phi}{:}f(\underline{a},\underline{a},a) \mathrel{\mid\!\mid\!\!\!\!\!\!\longrightarrow} f(b,b,a)$ extends the parallel step ϕ of Example 4.10, 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 4 right).

Guided by the example, we first 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 an extension $\bar{\phi}$ of ϕ_0, such that σ *itself* projects onto a cycle when projected over $\bar{\phi}$.

Consider repeatedly projecting the step ϕ in Example 4.10 over the cycle σ (see Figure 5). First, projecting ϕ over σ yields $\phi_0{:}f(a,\underline{a},a) \mathrel{\mid\!\mid\!\!\!\!\!\!\longrightarrow} f(a,b,a)$.

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

Figure 5. Repetition in computing residuals of ϕ after the cycle σ.

Next, projecting ϕ_0 over σ yields $\phi_1{:}f(\underline{a},a,a) \mathrel{\mid\!\mid\!\!\!\!\!\!\longrightarrow} f(b,a,a)$. Finally, projecting ϕ_1 over σ yields ϕ_0 again. That is, the series of residuals eventually becomes repetitive. This situation is completely general.

Lemma 4.13 (Repetition) *Let ϕ be a multi-step co-initial to a reduction cycle σ in an orthogonal PRS. Then there exists an extension ϕ_0 of ϕ, and a positive natural number n, such that $\phi_0/\sigma^n = \phi_0$.*

Proof Consider the infinite sequence of residuals ϕ/σ^i of ϕ after the i-fold repetition of σ, for arbitrary i. Since redex-patterns in a term in an orthogonal PRS do not have overlap, $\mathrel{\circ\!\!\!\!\!\longrightarrow}$ 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 $\sigma/(\phi/\sigma^i)$. Taking the first residual on the repetition yields the desired extension of ϕ. $\Box$

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 $f(a,a,a,a,a,a) \mathrel{+\!\!\!+\!\!\!\!\!\longrightarrow} f(b,a,a,b,a,a)$ 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 4.14 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 6 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

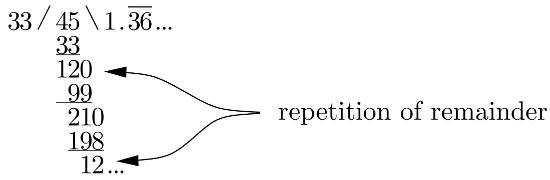

Figure 6. Repetition in computing decimal expansion when dividing by 33

as for residuals: since there are only finitely many distinct numbers modulo 33, the Pigeon Hole Principle yields that eventually a remainder, i.e. residual, reoccurs (in the example 12 reoccurs). Since the algorithm is deterministic, the decimal expansion is repetitive from there on.

Remark 4.15 The repetitive extension is monotonic in the sense that if σ is a cycle co-initial to both ϕ, ψ such that $\phi \subseteq \psi$, i.e. the redexes contracted by ϕ are a subset of those contracted by ψ, then ψ_0 extends ϕ_0.

Note that for such multi-steps ϕ and ψ, 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.

As a consequence of the Repetition Lemma, projecting *some repetition* σ^n of σs over the repetitive extension ϕ_0 of some multi-step ϕ, is a cycle again. But the goal was to find a multi-step $\bar{\phi}$ extending ϕ such that projecting σ

itself over $\bar{\phi}$ yields a cycle. To that end we *compress* all 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 5, showing a repetition of length two. For the residuals ϕ_0 and ϕ_1 on the repetition it holds:

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

Combining these two equalities yields, by distributivity of projection over union [14, Exercise 8.7.40(iv)] and commutativity of union:

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

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

Lemma 4.16 (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 $\bigcup_{0 \le i < n} \phi_0/\sigma^i$. Then $\bar{\phi}/\sigma = \bar{\phi}$ and $\bar{\phi}$ extends ϕ.*

Proof Using distributivity, we compute

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

That $\bar{\phi}$ extends ϕ follows from $\bar{\phi} = \bigcup_{0 \le i < n} \phi/\sigma^i \supseteq \phi/\sigma^0 = \phi$. $\square$

Remark 4.17 Although residuals of parallel steps are parallel steps again, 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 $a \to c(a)$, $c(x) \to x$, and $f(x, y) \to f(y, c(x))$, 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 4.18 Continuing Remark 4.14, we note that also compression can be applied to repeating decimals. To see this, consider computing the decimal expansion of $\frac{6}{7}$. Clearly, for each pair r, r' of consecutive remainders it holds:

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

Now consider the 'compression' $\bar{r} = \Pi_{0 \le i < n} r_i$ 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 $n = 6$.

Lemma 4.19 (Extension) *Let ϕ be a multi-step co-initial to a cycle σ. Then there exists an extension $\bar{\phi}$ of 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 $\bar{\phi}$ extends ϕ_0 for which it holds $\bar{\phi}/\sigma = \bar{\phi}$. Therefore, the other projection $\sigma/\bar{\phi}$ both starts and ends at the target of $\bar{\phi}$, so is a cycle. Since $\bar{\phi}$ extends ϕ_0 which extends ϕ, the result follows by transitivity. □

The extension $\bar{\phi}$ of ϕ used in the lemma is be called the *cyclic* extension of ϕ. The cyclic extension of ϕ of Example 4.10 is shown on the right in Figure 4.

Remark 4.20 Also the cyclic extension is monotonic. In fact, for a 'maximal' multi-step t^* from a term t, its repetitive and cyclic extensions coincide $t^*_0 = \bar{t}^*$.

Remark 4.21 Both the Repetition Lemma and the Compression Lemma can be entirely cast into the theory of finitely branching abstract residual systems [14, Sec. 8.7] for which union distributes over residuation. For instance, both lemmas hold for 'parallel moves' in the $\lambda\beta$-calculus.

4.4 Alternative proof of WHN ⇒ AC

Using the above results, an alternative proof of Theorem 2.6, i.e. of the fact that weakly head normalising orthogonal TRSs are acyclic, can be given.

For a proof by contradiction, assume there would exist a non-empty parallel cycle. Select among the parallel cycles of minimal length a head parallel cycle consisting of a head step ψ followed by a (possible empty) sequence τ of parallel steps. For any multi-step ϕ co-initial to ψ, the Extension Lemma yields an extension $\bar{\phi}$ of ϕ such that $\bar{\phi} = (\psi \cdot \tau)$ and $(\psi \cdot \tau)/\bar{\phi}$ is a parallel cycle again. We claim that it is in fact a head parallel cycle of the same length as $\psi \cdot \tau$. By repeated use of the claim we then reach the desired contradiction, since then also the target s of a reduction from t to head normal form, which exists by assumption, could be reduced further to a head redex, as shown in Figure 7.

To prove the claim note that by $\bar{\phi} = (\psi \cdot \tau)$ and the Non-Head Lemma, $(\psi \cdot \tau)/\bar{\phi}$ is a non-empty parallel cycle. By the minimality assumption the cycle has the same length as $\psi \cdot \tau$. In particular, the residual of ψ which is the first step on the cycle is non-empty, hence a head step by the Head Lemma.

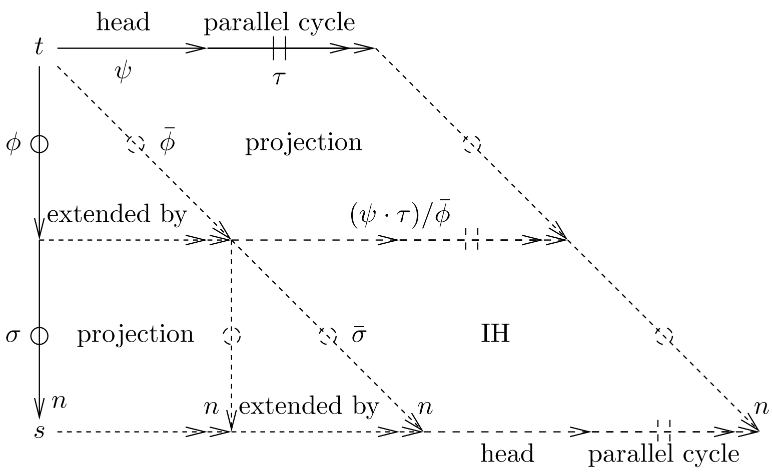

Figure 7. Projecting head parallel cycles inductively along extension

5 Acyclicity for λ-calculi

We consider acyclicity for various λ-calculi (with explicit substitutions).

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

Proof These are just Examples 2.13 and 3.4. □

Theorem 5.2 *For sub-calculi of the $\lambda\beta\eta$-calculus, WN $\Rightarrow$ AC.*

Proof It suffices to prove the Head Normalisation Theorem, i.e. that head strategies are hyper-normalising for $\lambda\beta\eta$-calculus. To that end, let t be a term which has a normal form. By η-postponement [3], 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 $\sigma_\eta{:}s \twoheadrightarrow u$. W.l.o.g. we may assume that s is in β-normal form.

We show that projecting such a reduction over a hyper-head reduction τ eventually results in an empty reduction. Thereto, we measure σ by first the number of η-redexes contracted in σ_η and second the μ-measure [12] of σ_β. We distinguish cases on whether the (first) step $\phi{:}t \to t'$ of τ projected over (use the weakly orthogonal projection of [12,14]) 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, and ϕ must be eliminated by σ_η since the latter was assumed to end in η-normal form. Thus the residual 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. □

Remark 5.3 We do not know whether the assumption of the theorem can be weakened from WN to WHN. One problem is that 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.

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

$$\begin{array}{rcl} (\lambda x.M)N & \to_{\mathrm{Beta}} & M[x{:=}N] \\ (M_1M_2)[x{:=}N] & \to_{@} & M_1[x{:=}N]M_2[x{:=}N] \\ (\lambda x.M)[y{:=}N] & \to_{\lambda} & \lambda x.M[y{:=}N] \\ x[x := N] & \to_{=} & N \\ y[x := N] & \to_{\neq} & y \end{array}$$

First, we consider the $\lambda\mathtt{x}^-$-calculus [4], which is the $\lambda\beta$-calculus extended with an explicit substitution operator $[:=]$ with rules as displayed, for $x \neq y$.

Theorem 5.4 *If $\lambda\mathtt{x}^-$ is weakly head normalising, then it is acyclic.*

Proof It suffices to adapt the Head Normalisation Theorem, proving that head reductions are hyper head normalising in $\lambda\mathtt{x}^-$. 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 is impossible by (the proof of) Theorem 3.3.

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 [12]. 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 [9]. 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, all 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. □

As for ordinary λ-calculi, this result in itself is not interesting as the $\lambda\mathtt{x}^-$-calculus is not weakly head normalising; it is interesting only 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, we define the sub-ARS of the $\lambda\beta$-calculus *induced* by $\mathcal{X}$ as $\to_\beta$ restricted to terms in $\mathcal{X}$, and verify that W(H)N is preserved.

Proposition 5.5 *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 $\mathcal{X}$.*

Hence, Theorem 5.4 extends to sub-calculi of $\lambda\mathtt{x}^-$.

Proof

- Note that any β-step on λ-terms may be simulated by a number of $\lambda\mathtt{x}^-$-steps [8, 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 [8, Lem. 5(1)] of t'. By [8, 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.

□

Finally, we consider the calculus of explicit substitutions $\lambda\sigma$ [1].

Theorem 5.6 *Weakly (head) normalising sub-calculi of $\lambda\sigma$ are acyclic.*

Proof This follows from [11, Thm. 3].

Here we present a direct proof that head strategies are hyper head normalising for $\lambda\sigma$, 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 8), 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

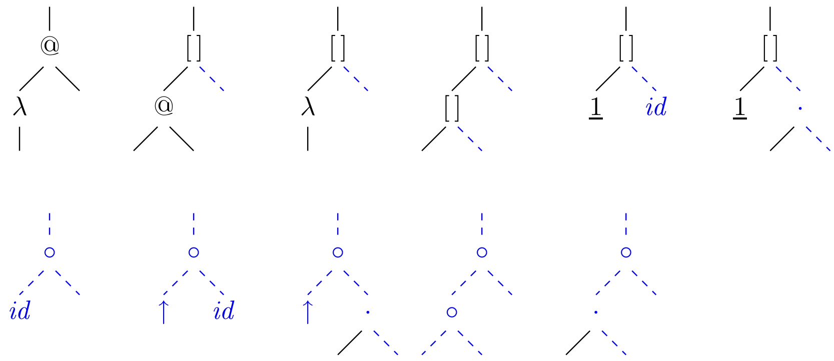

Figure 8. Left-hand sides of $\lambda\sigma$-rules, with *Substitution*s dashed/in blue

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 8).

Third, the explicit substitution calculus σ 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. $\square$

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 [10], display acyclic behaviour and hence unbounded growth; it *couldn't* have been otherwise! More precisely, we have:

Corollary 5.7 *Typed $\lambda\sigma$ is acyclic.*

Proof Let t be a typed $\lambda\sigma$-term and consider its family [3], i.e. the set all terms reachable from it w.r.t. the union of the reduction and the sub-term relations,[6] which is a sub-calculus. 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. $\square$

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 normal-

[6] Since 'sub-term steps' can be postponed, the same notion is obtained by considering only sub-terms of reachable terms.

ising, 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!

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,
- a theory of projections of cycles, and
- reduction of AC from λ-calculi with, to calculi without explicit substitution.

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, which we leave as conjectures:

- WHN $\Rightarrow$ AC, for weakly orthogonal PRSs.
- Head strategies are hyper-head-normalising for the $\lambda\beta\eta$-calculus.
- 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.

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] Baader, F. and T. Nipkow, "Term Rewriting and All That," Cambridge University Press, 1998.

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

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

[5] 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.

[6] Ketema, J., J. Klop and V. v. Oostrom, *Vicious circles in rewriting systems*, Report SEN-E0427, CWI (2004), Available on internet.

[7] 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.

[8] 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.

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

[10] 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.

[11] 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.

[12] Oostrom, V. 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.

[13] Oostrom, V. v., *Delimiting diagrams*, CKI Preprint 53, University of Utrecht (2004), Available on internet.

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