Infinitary Combinatory Reduction Systems*

Jeroen Ketema
Research Institute of Electrical Communication, Tohoku University
2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan
jketema@nue.riec.tohoku.ac.jp
and
Jakob Grue Simonsen
Department of Computer Science, University of Copenhagen (DIKU)
Universitetsparken 1, 2100 Copenhagen Ø, Denmark
simonsen@diku.dk

December 18, 2006

Abstract

Term rewriting is used for the modelling of computation in declarative languages and proof systems. Infinitary term rewriting is a class of models designed to model lazy languages with potentially infinite data structures, and infinitary logic. We define infinitary Combinatory Reduction Systems (iCRSs), thus providing the first notion of infinitary higher-order rewriting. The systems defined are sufficiently general that ordinary infinitary term rewriting and infinitary λ-calculus are special cases; using novel proof techniques, we generalise almost all fundamental results from those settings. Specifically, the following are proved:

1. Every reduction in a fully-extended, left-linear iCRS is compressible to a reduction of length at most ω.
2. Every complete development of an orthogonal set of redexes in an iCRS ends in the same term.
3. Every fully-extended, orthogonal iCRS is confluent modulo identification of hypercollapsing subterms.
4. Any outermost-fair, fair, or needed-fair strategy is normalising for the class of fully-extended, orthogonal iCRSs.

*Parts of this paper have previously appeared as [20] and [21].

Contents

1 Introduction

Term rewriting is a branch of mathematics and computer science that can generically model declarative programming, logic, universal algebra, and automated theorem proving. In term rewriting, equations are used as directed replacement rules where left-hand sides are replaced by right-hand sides, but not vice versa.

From a programming perspective, term rewriting affords easy modelling of function declaration and evaluation in declarative programming languages such as HASKELL, LISP, ML, and PROLOG. In these languages, functions are defined by equations, and a function call $\mathtt{foo}(\mathtt{a})$ is, conceptually, evaluated by replacing the call by the body of the function $\mathtt{foo}(x)$, actual parameter $\mathtt{a}$ being substituted for formal parameter x.

In, say, HASKELL, a simple function declaration to convert a list of integers to a list containing each of the integers multiplied by three would be:

$$\begin{aligned} \mathtt{treblelist}(n{:}ns) &= (n{*}3) : \mathtt{treblelist}(ns) \\ \mathtt{treblelist}([]) &= [] \end{aligned}$$

In a call $\mathtt{treblelist}(\mathtt{a})$ during computation — where we assume $\mathtt{a} \neq []$ — the call is matched against a pattern of the form $\mathtt{treblelist}(n{:}ns)$, and bindings are set up for the pattern variables n and ns. The pattern is then replaced by the right-hand side of the rule, i.e. the list $(n{*}3) : \mathtt{treblelist}(ns)$, and the computation continues. Thus, the function declaration above amounts to a set of two directed replacement rules, a *rewriting system*.

From a logical perspective, rewriting has its roots in equational logic where formulae are terms and equations describing relations between formulae give rise to rewrite rules; thus, proof systems (that may prove properties of programs or other objects) are naturally modelled using rewriting systems [3] — automated proof assistants such as Isabelle [28] and Coq [36] are based on term rewriting, e.g. in Coq proofs are terms and cut elimination is rewriting.

Our research in term rewriting is mostly motivated by programming; the remainder of the introduction thus focuses on that perspective.

Lazy Programming One particularly interesting feature of modern programming is the possibility to work explicitly with data structures that are *semantically* infinite — even though, in all concrete applications, program execution only examines a finite part of the data structure. For example, in HASKELL, the list $[\mathtt{0}..]$ denotes the *infinite* list of non-negative integers; indeed, the call $\mathtt{treblelist}([\mathtt{0}..])$ returns the infinite list of non-negative multiples of three. The infinite list is a first-class citizen and can be passed as an argument to a function, or can be returned as a result.

In HASKELL and other languages such lists make perfect sense due to *lazy evaluation*: The call $\mathtt{treblelist}([\mathtt{0}..])$ returns an infinite list, but does not actually *compute* any elements of it until program execution specifically asks for them. Programs may query as many elements as is possible on the concrete hardware the program is running on; moving the code to more powerful hardware will simply increase the number of elements of the list we can consider in a concrete computation.

If we choose the obvious route of modelling declarative languages by term rewriting, the presence of infinite lists generates subtle issues. The (first many) elements of $[\texttt{0}..]$ must somehow be evaluated in concrete programs; consider the following small code snippet:

$$\begin{aligned}\texttt{idl}(n{:}ns) &= n{:}\texttt{idl}(ns)\\ \texttt{idl}([]) &= []\end{aligned}$$

which is simply the identity function on lists. Now, the *result* of the call $\texttt{idl}([\texttt{0}..])$ must be some piece of data that represents an infinite list. Should the call $\texttt{idl}([\texttt{0}..])$ return the *term* $[\texttt{0}..]$ or something else? If the first option is chosen, what about functions f that are *semantically* equivalent to `idl`, but might have much more involved definitions? Rice's Theorem prevents us from determining whether $f(x) = \texttt{idl}(x)$ for all x, so we cannot always 'guess' if $f([\texttt{0}..])$ should return $[\texttt{0}..]$. Hence, a system of rules for modelling such computation cannot in general hope to have anything remotely resembling $[\texttt{0}..]$ as the output value of $f([\texttt{0}..])$.

We could go for the second option and simply return whatever (infinite) data structure the computation of $f([\texttt{0}..])$ would build when given sufficient time. This data structure could be a wrapped 'black box' that could be passed around to other functions that could use it to generate as many elements as were needed. This is a perfectly viable solution, but still leaves the question open of how to compare such infinite data structures. For instance, how would we prove that the black box actually equals $\texttt{idl}([\texttt{0}..])$? A possible solution is simply to let the return value of $f([\texttt{0}..])$ be the *infinite* term that computation of the call would yield if it were allowed to go on indefinitely, i.e. $[\texttt{0}, \texttt{1}, \texttt{2}, \ldots]$.

We are then faced with the problem of how to succinctly represent such infinite terms, how to ensure that computation — application of rewrite rules — achieves progressively better approximations to the final data structure as the number of steps in the computation increases to infinity. These problems are the motivation behind research in infinitary term rewriting.

Infinitary Rewriting Standard techniques in term rewriting of finite terms can be used to prove interesting properties of declarative programs, e.g. termination; but these techniques now require serious rethinking: A program terminating on *finite* lists may fail to terminate on *infinite* ones (e.g. a program iterating over each element of the list).

Alas, most of the truly useful theory of ordinary term rewriting of finite terms turns out, unfortunately, to be inapplicable when infinite terms are involved.

An easy and powerful way around many of these problems is to define a notion of *convergence* of program execution: Each computation step can be viewed as an element in a sequence of terms; the sequence may be infinite, but it will *converge* to some well-defined result, e.g. an infinite list.

If the notion of convergence is defined properly, proving a property of an infinite sequence can then have implications for *finite* program executions as well. This is the reason for introducing the notion of so-called *strong convergence*, defined by Kennaway et al. in their landmark paper [18]. Strong convergence employs both the ordinary theory of metric spaces and a special 'syntactic property' on rewrite steps. The intuition is that idealised 'infinite' computations should never touch the 'top' part of a term, e.g. the finite prefixes of an infinite list after some specific *finite* number of computation steps have been performed: Thus, greater and greater initial parts of the 'result' of a program will be completed in *finite* time. If these successively larger parts then converge in some appropriate metric on terms, they are said to *converge strongly* to their limit — a possibly infinite term that may be viewed as the result of the — idealised — infinite computation. Conversely, if we shut down the computation after a finite number of steps, strong convergence ensures that we *will* have a good approximation to the final 'ideal' result.

Example 1.1. The computation of `treblelist`([0..]) transforms arbitrarily long finite prefixes of [0..] into lists with their elements multiplied by three. We can make this statement more formal by showing that the sequence of rewrite steps that models the computation of `treblelist`([0..]) generates a sequence of terms that converges strongly to the list $[0, 3, 6, 9, \ldots]$.

Strong convergence requires a metric in the ordinary sense; such a metric is given by letting, two (finite or infinite) lists have distance 2^{-k} from each other if they have the same first $k - 1$ elements, but their kth elements differ from each other.

In the above we have:

$$\begin{aligned}\texttt{treblelist}([0..]) &\to \texttt{0} : \texttt{treblelist}([1..]) \\ &\to \texttt{0} : \texttt{3} : \texttt{treblelist}([2..]) \\ &\to \texttt{0} : \texttt{3} : \texttt{6} : \texttt{treblelist}([3..]) \\ &\to \cdots\end{aligned}$$

As $[0, 3, 6, 9, ..]$ is syntactic sugar for $0 : 3 : 6 : 9 : \ldots$, the nth element in the sequence of terms in the computation above is at a distance of 2^{-n} from the infinite term $[0, 3, 6, 9, ...]$, i.e. the distance between two successive computation steps converges to 0 as the number of computation steps goes to infinity. In addition, no computation steps are performed in the initial part of the result (the part of the list containing the first n elements) after n steps have been performed — the steps occur 'deeper' in the terms — i.e. we can extract meaningful partial results after finite time has elapsed.

Higher-Order Functions The theory of first-order programming with potentially infinite data structures has been developed successfully using infinitary rewriting since the mid-nineties. However, first-order constructs are insufficient for the modern programmer — his arsenal includes *higher-order* functions (functions taking functions as arguments) that cannot be modelled by the first-order constructs of term rewriting without lengthy and unintuitive encodings. For instance, the function map that applies a function to each element of a list in succession:

$$\texttt{map}\, f\, (x{:}xs) = (f\, x) : \texttt{map}\, f\, xs$$
$$\texttt{map}\, f\, [] = []$$

The classical 'theorist's approach' to handling higher-order functions such as map is to simply appeal to the machinery of λ-*calculus* [4]. Function evaluation in λ-calculus is expressed through its single rewrite rule

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

where $M[x := N]$ is the substitution of the parameter N for the free occurrences of variable x in the function body M. The extension of λ-calculus to infinite terms and computations [16] affords an idealised model of function evaluation, including higher-order function evaluation that can handle constructs such as map, but it is quite awkward to take a real-world functional program and encode it directly in λ-calculus.

A much more straightforward encoding is possible by using one of the variants of so-called *higher-order rewriting* [1, 22, 24, 29, 40, 12]. For instance, in the syntax of one of these variants — *Combinatory Reduction Systems* (CRSs) — the definition of map becomes:

$$\texttt{map}([z]F(z), \texttt{cons}(X, XS)) \to \texttt{cons}(F(X), \texttt{map}([z]F(z), XS))$$
$$\texttt{map}([z]F(z), \texttt{nil}) \to \texttt{nil}$$

i.e. just a de-sugared version of the declaration of map where variable bindings have been made explicit.

It is important to note that although the syntax of the different forms of higher-order rewriting varies, the forms are just more-or-less equivalent ways of easing the burden on the programming language designer: If he is willing to spend a few minutes expressing his syntax as a higher-order rewriting system of some form, the entire theory developed for that particular form can be brought to bear on his problems. For instance, if his syntax satisfies a few easily checkable conditions — called 'orthogonality' in the rewriting vernacular — he can appeal to standard results to show a number of results related to program behaviour. For instance, he can show that his language will be deterministic — by appealing to the so-called 'confluence property'. By appealing to other standard techniques, he can show that there will be execution strategies for any program in his language that always will, if at all possible, get around all ways of going into infinite loops and actually yield a well-defined result — called 'normalisation'. All of these desiderata will be available to the language designer *without the onus of him having to prove it using the specifics of his own invention*[1].

[1] Indeed, this does away with the '...prove the Church-Rosser property by the method of Tait and Martin-Löf in the style of Aczel (see appendix)'-sentence found in many papers on declarative programming.

Infinitary rewriting and higher-order functions Unfortunately, higher-order functions like `map` cannot be treated by the machinery developed in infinitary rewriting so far. Graph rewriting [31] affords a way of treating such functions in the setting of lazy programming, but (higher-order) graph rewriting does not yet have the same array of generally applicable results that can be brought to bear as does traditional term rewriting. Thus, a true extension of infinitary rewriting to the higher-order setting should be defined and as many of the 'usual' results as possible should be re-derived.

This paper is devoted to exactly that. We define *infinitary* Combinatory Reduction Systems and thus extend the modelling of lazy declarative programming using infinitary term rewriting to the higher-order case. In doing so, we generalise most of the results known to hold for infinitary (first-order) term rewriting and for infinitary λ-calculus.

We note that as rewriting is traditionally also used for computing with equational logic, our work also allows for modelling of formulae in *infinitary logic* with quantifiers and bound variables [26] in the same fashion as is usually done in ordinary rewriting with ordinary logic; we do not yet know whether this has any implications for the possible use of infinitary logic in any practical matters.

1.1 Brief History of Infinitary Rewriting

The approach of considering infinite terms by metric spaces was originally pioneered by Arnold and Nivat [2]; alternative approaches considered defining infinite terms by means of partial functions [13, 6].

Rewriting of infinite terms was first considered in the context of first-order rewriting where rewriting systems were equipped with a very liberal notion of potentially infinite reduction, called *weak* or *Cauchy* convergence, by Dershowitz, Kaplan, and Plaisted [7, 8]. Upon discovering subtle problems with the approach, the authors published a final and corrected version of their results [9], but many of the standard results from ordinary term rewriting could not be recovered.

To alleviate this, Kennaway, Klop, Sleep, and De Vries, inspired by Farmer and Watro's paper [11], considered a more restrictive notion of infinite reduction, based on so-called *strong convergence*, that has become the de facto standard in infinitary rewriting; their results were published in the landmark paper [18] that pinned down basic results related to confluence and normalisation for first-order term rewriting systems with potentially infinite terms and reductions. The first step towards higher-order infinitary rewriting was taken by the same authors when they considered infinitary λ-calculus [16].

The field has since expanded, leading to consideration of so-called *meaninglessness* (identifying a class of terms essentially having no 'good' definable semantics) [19], observational equivalence [10], alternative approaches to defining infinite terms and their accompanying rewrite relation [5], modular properties [34], and uniform normalisation [23].

1.2 Overview of Present Paper

We define and prove basic properties of infinitary Combinatory Reduction Systems. The technical development follows mostly the tried-and-true technique in rewriting of establishing properties of *developments* of sets of redexes in so-called *orthogonal* systems. We combine techniques from infinitary rewriting with methods for proving reduction strategies normalising in ordinary (finitary) rewriting to prove *confluence modulo* results for orthogonal iCRSs and to prove that a number of strategies are normalising for such systems.

The main contributions of the paper are threefold:

- The introduction of several new techniques for proving results in infinitary rewriting that go beyond the extant tools so far seen in the field; in particular, the use of *essentiality* — see Section 8 — is worthy of mention. These techniques are substantially different from prior known techniques used in first-order infinitary rewriting and might possibly have uses elsewhere.
- An extension of the confluence (modulo) results so far known in infinitary rewriting to the general higher-order setting. In particular, our results generalise those already known to hold for (first-order) infinitary rewriting and infinitary λ-calculus[2].
- An extension of the normalisation results of infinitary rewriting by using the technique of essentiality to show that needed-fair, fair, and outermost-fair reductions are normalising.

It should be noted that Lisper has defined a separate notion of infinitary Combinatory Reduction Systems in [25] and proves a number of preliminary results for these. His notion of infinite terms is essentially an instance of ours: It contains only rules with finite right-hand sides, and many of the results concerning, e.g. compression, impose further restrictions on the systems considered. The restrictions materialise in several crucial places, e.g. when unfoldings for higher-order rules are considered, Lisper recommends switching to a *first-order* combinator system.

Structure of the paper The specific layout of the paper is as follows: Section 2 contains preliminary definitions. Sections 3 and 4 introduce terms and rewriting. Section 5 proves that every well-behaved rewrite sequence of transfinite length can be 'compressed' to one of length at most ω. Sections 6, 7 and 8 set up fundamental properties of *orthogonal* systems and the main tools in such systems that will be used to prove later properties; Section 8 in particular introduces the technique of *essentiality* which is a fulcrum of both the confluence and the normalisation results to follow. Section 9 expounds one of the

[2] It should be noted that we extend our preliminary findings in [21]: There, a confluence modulo result was proven for systems where every rule had a finite right-hand side. In the present paper, we strengthen that result by allowing infinite right-hand sides.

main results of the paper: That fully-extended, orthogonal systems are confluent. Section 10 shows that a variety of reduction strategies are normalising for fully-extended, orthogonal systems. Section 11 concludes and affords pointers for further work. Appendix A contains the — lengthy — proofs of two key ancillary results: Proposition 6.6 and Lemma 6.7

1.3 Bluffer's Guide

Readers with prior knowledge of rewriting and the syntax of (ordinary, finite) Combinatory Reduction Systems can make do with noting that meta-terms are simply formed by interpreting the rules for meta-term formation coinductively. A serious caveat is that 'infinite chains of immediately nested meta-variables' must be avoided — see Section 3.4.

For the reader with prior knowledge of infinitary rewriting, we introduce metrics on terms and transfinite reductions in the usual manner; compression requires a more substantial analysis than usual due to the fact that nestings can occur in reduction steps — see Section 5. Due to the problem of redexes appearing after an infinite number of steps because of variables being 'pushed out' of a term, we are generally forced to require that all rules are fully-extended — see Definition 4.5.

The main properties of confluence and normalisation are proved by going through the usual route of giving results for developments of sets of redexes in orthogonal systems — Section 6. These results hinge upon a version of the notion of *paths* found in [17], but the related results require a much more involved treatment in the higher-order setting. Confluence and normalisation are proved by appealing to a certain technique — *essentiality* — used for proving normalisation in finitary rewriting, heavily massaged to be applicable here. *Paths* are once again needed to ensure that results are applicable to systems with rules having infinite right-hand sides; this again forces us to construct proofs markedly different from those of the first-order setting.

2 Preliminaries

Prior knowledge of CRSs [22, 24, 40] and infinitary rewriting [17] is not required, but will greatly improve the reader's understanding of the text. Throughout, infinitary Term Rewriting Systems are invariably abbreviated as iTRSs and infinitary λ-calculus is abbreviated as iλc.

We assume a signature Σ, each element of which has finite arity. We also assume a countably infinite set of variables and, for each finite arity, a countably infinite set of meta-variables. Countably infinite sets are sufficient, given that we can employ 'Hilbert hotel'-style renaming. We denote the first infinite ordinal by ω, and arbitrary ordinals by $\alpha, \beta, \gamma, \ldots$.

The set of *finite meta-terms* is defined as follows:

1. each variable x is a finite meta-term,

2. if x is a variable and s is a finite meta-term, then $[x]s$ is a finite meta-term,

3. if Z is a meta-variable of arity n and $s_1, \ldots, s_n$ are finite meta-terms, then $Z(s_1, \ldots, s_n)$ is a finite meta-term,

4. if $f \in \Sigma$ has arity n and $s_1, \ldots, s_n$ are finite meta-terms, then $f(s_1, \ldots, s_n)$ is a finite meta-term.

A finite meta-term of the form $[x]s$ is called an *abstraction*. Each occurrence of the variable x in s is *bound* in $[x]s$, and each subterm of s is said to occur in the *scope* of the abstraction. If s is a finite meta-term, we denote by $root(s)$ the root symbol of s.

The *set of positions* of a finite meta-term s, denoted $\mathcal{P}os(s)$, is the set of finite strings over $\mathbb{N}$, with ϵ the empty string, such that:

1. if $s = x$ for some variable x, then $\mathcal{P}os(s) = \{\epsilon\}$,

2. if $s = [x]t$, then $\mathcal{P}os(s) = \{\epsilon\} \cup \{0 \cdot p \mid p \in \mathcal{P}os(t)\}$,

3. if $s = Z(t_1, \ldots, t_n)$, then $\mathcal{P}os(s) = \{\epsilon\} \cup \{i \cdot p \mid 1 \leq i \leq n, p \in \mathcal{P}os(t_i)\}$,

4. if $s = f(t_1, \ldots, t_n)$, then $\mathcal{P}os(s) = \{\epsilon\} \cup \{i \cdot p \mid 1 \leq i \leq n, p \in \mathcal{P}os(t_i)\}$.

The *depth* of a position p, denoted $|p|$, is the number of characters in p. Given $p, q \in \mathcal{P}os(s)$, we write $p \leq q$ and say that p is a *prefix* of q and that q is a *suffix* of p, if there exists an $r \in \mathcal{P}os(s)$ such that $p \cdot r = q$. If $r \neq \epsilon$, we also write $p < q$ and say that the prefix is *strict*. Moreover, if neither $p \leq q$ nor $q \leq p$, we say that p and q are *parallel*, which we write as $p \parallel q$.

We denote by $s|_p$ the subterm of s that occurs *at* position $p \in \mathcal{P}os(s)$. Moreover, if $q \in \mathcal{P}os(s)$ and $p < q$, we say that the subterm at position p occurs *above* q. Finally, if $p > q$, then we say that the subterm occurs *below* q.

3 Terms and Substitutions

We now proceed to define the main objects of study, namely *meta-terms* and *terms*. Furthermore, we define substitutions on terms which will be crucial in defining the rewrite relation on terms.

As it turns out, the most straightforward and liberal definition of meta-terms has rather poor properties: Applying a substitution need not necessarily yield a well-defined term. Therefore, we also introduce an important restriction on meta-terms: the *finite chains property*. This property will also prove crucial in obtaining positive results later in the paper.

3.1 Meta-Terms and Terms

In iTRSs and iλc, terms are defined by introducing a metric over the set of finite terms and taking the completion of that metric. That is, taking the least set of objects (with respect to set inclusion) containing the set of finite terms such that every Cauchy sequence converges [2, 18, 16] — this set will contain both the finite and infinite terms. Intuitively, with respect to such a metric, two terms s and t are close to each other if the first 'conflict' between them occurs at great depth. In iTRSs, a conflict is a position p such that $root(s|_p) \neq root(t|_p)$. In i$\lambda$c, a conflict is defined similarly, but also takes into account α-equivalence. The metric, denoted $d(s,t)$, is defined as 0 if no conflict occurs between s and t and is otherwise defined as 2^{-k}, where k denotes the minimal depth such that a conflict occurs between s and t. We take a similar approach in this paper.

To define terms and meta-terms for iCRSs, we first define the notions of a conflict and α-equivalence for *finite* meta-terms. In the definition, $s[x \to y]$ denotes the replacement in s of the occurrences of the free variable x by the variable y.

Definition 3.1. Let s and t be finite meta-terms. A *conflict* of s and t is a position $p \in \mathcal{P}os(s) \cap \mathcal{P}os(t)$ such that:

1. if $p = \epsilon$, then $root(s) \neq root(t)$,
2. if $p = i \cdot q$ for $i \geq 1$, then $root(s) = root(t)$ and q a conflict of $s|_i$ and $t|_i$,
3. if $p = 0 \cdot q$, then $s = [x_1]s'$ and $t = [x_2]t'$ and q a conflict of $s'[x_1 \to y]$ and $t'[x_2 \to y]$, where y does not occur in either s' or t'.

The finite meta-terms s and t are *α-equivalent* if no conflict exists between them.

The metric is now defined precisely as in the case of iTRSs and iλc:

Definition 3.2. The metric d on the set of finite meta-terms is defined as follows:

$$d(s,t) = \begin{cases} 0 & \text{if } s \text{ and } t \text{ are } \alpha\text{-equivalent} \\ 2^{-k} & \text{otherwise} \end{cases}$$

where k is the minimal depth such that a conflict occurs between s and t.

Example 3.3. The meta-terms $s = [x]Z(x, f(x))$ and $t = [y]Z(y, f(y))$ satisfy $d(s,t) = 0$. Moreover, if $t' = [y]Z(y, f(z))$, then the only conflict between s and t' occurs at position 021 and, hence, $d(s,t') = 2^{-3} = \frac{1}{8}$.

Precisely following the definition of terms in the case of iTRSs and iλc, we define the set of meta-terms.

Definition 3.4. The set of *meta-terms* is the metric completion of the set of finite meta-terms with respect to the metric d.

By definition of metric completion, the set of finite meta-terms is a subset of the set of meta-terms. Moreover, we can uniquely extend the metric d to a metric on the set of meta-terms, which we also denote by d.

Example 3.5. Any finite meta term, e.g. $[x]Z(x, f(x))$ is a meta-term. Moreover, $Z(Z(Z(\ldots)$ is a meta-term, as is $Z_1([x_1]x_1, Z_2([x_2]x_2, \ldots))$.

The notions of a set of positions and a subterm of a finite meta-term carry over directly to the meta-terms, we use the same notation in both cases.

The set of terms can now be defined like in the finite case [22, 24, 40], i.e. by barring meta-variables from occurring. The only difference is that meta-terms now occur in the definition instead of finite meta-terms.

Definition 3.6. The set of *terms* is the largest subset of the set of meta-terms such that no meta-variables occur in the meta-terms.

Both the set of (infinite) first-order terms and the set of (infinite) λ-terms are easily shown to be included in the set of terms.

The definition of context carries over directly from the finite case:

Definition 3.7. A *context* is a term over $\Sigma \cup \{\Box\}$ where $\Box$ is a fresh nullary function symbol. A *one-hole* context is a context in which precisely one $\Box$ occurs.

Henceforth, we use $f^n(s)$ for any $n \in \mathbb{N}$ and term s to denote the following inductively defined term:

$$f^n(s) = \begin{cases} s & \text{if } n = 0 \\ f(f^m(s)) & \text{if } n = m + 1 \end{cases}$$

Moreover, we use f^ω to denote the term that is the solution of the recursive equation $s = f(s)$ or more informally $f(f(\ldots f(\ldots)))$.

As mentioned in the introduction to this section, we shall later define a restriction on meta-terms called the *finite chains property*. Intuitively, a *chain* is a sequence of contexts in a meta-term that occur 'nested right below each other'. Formally:

Definition 3.8. Let s be a meta-term. A *chain* in s is a sequence of (context, position)-pairs $(C_i[\Box], p_i)_{i<\alpha}$, with $\alpha \leq \omega$, such that for each $(C_i[\Box], p_i)$ there exists a term t_i with $C_i[t_i] = s|_{p_i}$ and $p_{i+1} = p_i \cdot q$ where q is the position of the hole in $C_i[\Box]$. If $\alpha < \omega$, respectively $\alpha = \omega$, then the chain is called *finite*, respectively *infinite*.

3.2 Alternative Definition of Meta-Terms and Terms

Above, we followed the beaten path of defining the set of meta-terms as the metric completion of the set of finite meta-terms *where finite terms were considered equal up to α-equivalence*. An alternative, and equally reasonable, approach is to use ordinary syntactic equality on finite meta-terms, to take the metric completion of the set, and to subsequently define meta-terms as α-equivalence classes in the resulting set.

To cover as much ground as possible, we now proceed to show that the latter method is equivalent to the former, in the obvious technical sense. This requires a massaged version of Definition 3.1:

Definition 3.9. Let s and t be finite meta-terms. A *raw conflict* of s and t is a position $p \in \mathcal{P}os(s) \cap \mathcal{P}os(t)$ such that:

1. if $p = \epsilon$, then $root(s) \neq root(t)$,
2. if $p = i \cdot q$, then $root(s) = root(t)$ and q a conflict of $s|_i$ and $t|_i$.

Hence, a raw conflict is a conflict in which differences in the encountered abstractions are considered to be real differences.

Example 3.10. The meta-terms $s = [x]Z(x, f(x))$ and $[y]Z(y, f(y))$ have a raw conflict at position ϵ. Moreover, s has a raw conflict with $[x]Z(x, f(y))$ at position 021.

Definition 3.11. The *raw metric* d_r on the set of finite meta-terms is defined as follows:

$$d_r(s,t) = \begin{cases} 0 & \text{if } s = t \\ 2^{-k} & \text{otherwise} \end{cases}$$

where k is the minimal depth such that a raw conflict occurs between s and t.

Definition 3.12. The set of *raw pre-meta-terms* is the metric completion of the set of finite meta-terms with respect to the raw metric d_r.

As before, the notion of position carries over directly to raw pre-meta-terms.

Example 3.13. We have that $Z([x]Z'(x, y, Z([x]Z'(x, y, \ldots))))$ is a raw pre-meta-term, as are $Z(Z(Z(\ldots)))$ and $[x]Z(x)$.

The definition of α-equivalence for finite meta-terms carries over straightforwardly to the set of raw meta-terms, except that coinduction is now involved:

Definition 3.14. Let s be a raw pre-meta-term and x and y variables. Then, $s[x \to y]$ denotes the raw pre-meta-term obtained by replacing in s all of the occurrences of the free variable x by the variable y, defined coinductively by:

1. $x[x \to y] = y$,
2. $z[x \to y] = z$, if $x \neq z$,
3. $([x]s)[x \to y] = [x]s$,
4. $([z]s)[x \to y] = [z](s[x \to y])$, if $x \neq z$,
5. $(Z(s_1, \ldots, s_n))[x \to y] = Z(s_1[x \to y], \ldots, s_n[x \to y])$,
6. $(f(s_1, \ldots, s_n))[x \to y] = f(s_1[x \to y], \ldots, s_n[x \to y])$.

Example 3.15. We have

$$(f([x]Z'(x, y, \ldots)))[y \to z] = f([x]Z'(x, z, \ldots))$$

and

$$(f([x]Z'(x, y, \ldots)))[x \to z] = f([x]Z'(x, y, \ldots))\,.$$

Given the above definition, Definition 3.1 carries over directly to raw pre-meta-terms. Hence, we can define:

Definition 3.16. The raw pre-meta-terms s and t are *α-equivalent*, denoted $s =_\alpha t$, if no conflict between s and t exists.

It is routine to verify that $=_\alpha$ is an equivalence relation on the set of raw meta-terms.

Definition 3.17. The set of *raw meta-terms* is the set of equivalence classes over the set of raw pre-meta-terms with respect to $=_\alpha$.

The set of *raw terms* is the largest subset of the set of raw meta-terms containing only equivalence classes having a representative without meta-variables.

It is straightforward to see that if $[S]$ is a raw meta-term and $s, t \in [S]$ are representatives of $[S]$, then s contains a meta-variable iff t contains one. The definition of raw term is thus robust.

We need to ensure that the two different ways of constructing infinite meta-terms yield identical behaviour with respect to the metric d. This is the contents of the following theorem:

Theorem 3.18. *Let $[S]$ and $[T]$ be distinct raw meta-terms and let $s, s' \in [S]$ and $t, t' \in [T]$, then $d(s,t) = d(s',t')$.*

Proof. By symmetry, it suffices to prove that $d(s,t) = d(s,t')$. As $t, t' \in [T]$, we have $\mathcal{P}os(t) = \mathcal{P}os(t')$. Moreover, as $[S]$ and $[T]$ are distinct, a conflict occurs between s and t at some position p. Without loss of generality we may assume that p is a position of minimal depth with this property.

We proceed by induction on p:

- If $p = \epsilon$, then there is clearly a conflict between s and t' at position ϵ, otherwise t and t' would not be α-equivalent.
- If $p = i \cdot q$ for $i \geq 1$, then $root(s) = root(t) = root(t')$. By the induction hypothesis, we have $d(s|_i, t|_i) = d(s|_i, t'|_i) = 2^{-|q|}$, and the result follows.
- If $p = 0 \cdot q$, then $s =_\alpha [x]\tilde{s}$, $t =_\alpha [x]\tilde{t}$, and $t' =_\alpha [x]\tilde{t'}$ for suitable meta-terms $\tilde{s}$, $\tilde{t}$, and $\tilde{t'}$ and a variable x. The induction hypothesis yields $d(\tilde{s}, \tilde{t}) = d(\tilde{s}, \tilde{t'}) = 2^{-|q|}$, and the result follows. □

The above theorem ensures that d induces a well-defined metric on the set raw meta-terms and we could have chosen the set as an alternative for the set of meta-terms (in practice always working with some representative of each equivalence class). Thus, both 'natural' ways of defining infinite terms modulo α-equivalence yield, for all purposes, sets of infinite terms with identical behaviour. We find the definition of infinite terms introduced in the previous subsection the more natural one to work with.

3.3 Substitutions

We next define substitutions. The required definitions are the same as in the case of CRSs [24, 40], except that the interpretation of the definition is coinductive (due to the presence of infinite terms and meta-terms), rather than inductive. This is identical to what is done in the case of iTRSs and iλc in relation to the finite systems on which those systems are based. Below, we use $\vec{x}$ and $\vec{t}$ as short-hands for, respectively, the sequences $x_1, \ldots, x_n$ and $t_1, \ldots, t_n$ with $n \geq 0$. Moreover, we assume n fixed in the next two definitions.

Definition 3.19. A *substitution* of the terms $\vec{t}$ for distinct variables $\vec{x}$ in a term s, denoted $s[\vec{x} := \vec{t}]$, is defined as:

1. $x_i[\vec{x} := \vec{t}] = t_i$,
2. $y[\vec{x} := \vec{t}] = y$, if y does not occur in $\vec{x}$,
3. $([y]s')[\vec{x} := \vec{t}] = [y](s'[\vec{x} := \vec{t}])$,
4. $f(s_1, \ldots, s_m)[\vec{x} := \vec{t}] = f(s_1[\vec{x} := \vec{t}], \ldots, s_m[\vec{x} := \vec{t}])$.

The above definition implicitly takes into account the usual variable convention [4] in the third clause to avoid the binding of free variables by the abstraction. We now define substitutes (adopting this name from Kahrs [15]).

Definition 3.20. An *n-ary substitute* is a mapping denoted $\underline{\lambda} x_1, \ldots, x_n.s$ or $\underline{\lambda}\vec{x}.s$, with s a term, such that:

$$(\underline{\lambda}\vec{x}.s)(t_1, \ldots, t_n) = s[\vec{x} := \vec{t}]\,. \tag{1}$$

Reading Equation (1) from left to right yields a rewrite rule:

$$(\underline{\lambda}\vec{x}.s)(t_1, \ldots, t_n) \to s[\vec{x} := \vec{t}]\,.$$

The rule can be seen as a *parallel β-rule*. That is, a variant of the β-rule from iλc which simultaneously substitutes multiple variables. We call the root of $(\underline{\lambda}\vec{x}.s)$ the $\underline{\lambda}$-abstraction and the root of the left-hand side of the parallel β-rule the $\underline{\lambda}$-application.

Definition 3.21. Let σ be a function that maps meta-variables to substitutes such that, for all $n \in \mathbb{N}$, if Z has arity n, then so does $\sigma(Z)$.

A *valuation* induced by σ is a map $\bar{\sigma}$ from meta-terms to terms that has the following properties:

1. $\bar{\sigma}(x) = x$,
2. $\bar{\sigma}([x]s) = [x](\bar{\sigma}(s))$,
3. $\bar{\sigma}(Z(s_1, \ldots, s_m)) = \sigma(Z)(\bar{\sigma}(s_1), \ldots, \bar{\sigma}(s_m))$,
4. $\bar{\sigma}(f(s_1, \ldots, s_m)) = f(\bar{\sigma}(s_1), \ldots, \bar{\sigma}(s_m))$.

Similar to Definition 3.19, the above definition implicitly takes into account the variable convention in the second clause to avoid the binding of free variables by the abstraction.

From an operational point-of-view the definition of a valuation yields a straightforward two-step way of applying it to a meta-term: In the first step each subterm of the form $Z(t_1, \ldots, t_n)$ is replaced by a subterm of the form $(\underline{\lambda}\vec{x}.s)(t_1, \ldots, t_n)$. In the second step Equation (1) is applied to each subterm of the form $(\underline{\lambda}\vec{x}.s)(t_1, \ldots, t_n)$ as introduced in the first step.

In view of the rewrite rule introduced immediately below Definition 3.20 the second step can be seen as a complete development of the parallel β-redexes introduced in the first step. This is obviously a complete development in a variant of iλc. The only rule of the variant is the parallel β-rule and the signature of the variant contains, besides the $\underline{\lambda}$-application and the $\underline{\lambda}$-abstraction, the abstractions, the meta-variables, and the elements of Σ of the considered iCRS.

In the finite case [22, Remark II.1.10.1], the application of a valuation to a meta-term yields a unique term, i.e. *valuations are always well-defined*. Unfortunately, this is no longer the case when infinite meta-terms are considered:

Example 3.22. Consider the meta-term

$$Z(Z(\ldots Z(\ldots)))\,.$$

Applying the valuation induced by any map that satisfies $Z \mapsto \underline{\lambda}x.x$ yields:

$$(\underline{\lambda}x.x)((\underline{\lambda}x.x)(\ldots(\underline{\lambda}x.x)(\ldots)))$$

This '$\underline{\lambda}$-term' has no complete development, as no matter how many parallel β-redexes are contracted, it reduces only to itself and not to a term.

The previous problem does not depend one unique meta-variable being present in the meta-term. The same behaviour can occur with different meta-variables of different arities if we can define a valuation that assigns $\underline{\lambda}\vec{x}.y$ to each meta-variable Z in the meta-term with y in $\vec{x}$ such that y corresponds to an argument of Z which is a meta-variable.

In the above example, $Z(Z(\ldots Z(\ldots)))$ is not in the domain of $\bar{\sigma}$ due to the fact that *no* function can have the properties needed to be a valuation induced by σ *and* be defined on $Z(Z(\ldots Z(\ldots)))$.

The problem is even more intricate: The action of applying a valuation by reducing in the corresponding '$\underline{\lambda}$-term' is not necessarily confluent.

Example 3.23. Consider a signature with nullary functions symbols a and b. Moreover, consider the meta-term

$$Z(a, Z(b, Z(a, Z(b, Z(\ldots)))))\,.$$

Applying the valuation that assigns to Z the substitute $\underline{\lambda}xy.y$ yields the '$\underline{\lambda}$-term':

$$(\underline{\lambda}xy.y)(a, (\underline{\lambda}xy.y)(b, (\underline{\lambda}xy.y)(a, (\underline{\lambda}xy.y)(b, (\underline{\lambda}xy.y)(\ldots)))))\,,$$

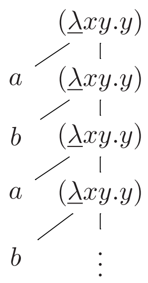


Figure 1:

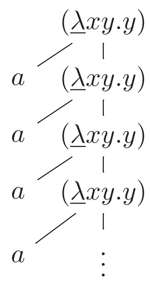


Figure 2:

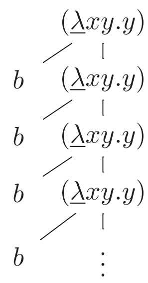


Figure 3:

which is depicted in Figure 1. The term reduces by means of two different developments to the '$\underline{\lambda}$-terms':

$$(\underline{\lambda}xy.y)(a, (\underline{\lambda}xy.y)(a, (\underline{\lambda}xy.y)(a, (\underline{\lambda}xy.y)(a, (\underline{\lambda}xy.y)(\ldots))))) ,$$

as depicted in Figure 2, and:

$$(\underline{\lambda}xy.y)(b, (\underline{\lambda}xy.y)(b, (\underline{\lambda}xy.y)(b, (\underline{\lambda}xy.y)(b, (\underline{\lambda}xy.y)(\ldots))))) ,$$

as depicted in Figure 3. These last two '$\underline{\lambda}$-terms' have no common reduct with respect to parallel β-reduction. They reduce only to themselves. Note that this problem also occurs in iλc [16, Section 4].

The situation is thus unsatisfactory: We would like valuations to be defined on as many meta-terms as possible. By the above examples, this is not possible in general. We thus need to identify a class of meta-terms that avoids these problems, yet is sufficiently expressive.

3.4 Finite Chains Property

The examples exhibiting problems with valuations all share a common feature: They involve meta-terms with infinite chains of meta-variables. Formally:

Definition 3.24. Let s be a meta-term. A *chain of meta-variables* in s is a chain $(C_i[\Box], p_i)_{i<\alpha}$ in s, with $\alpha \leq \omega$, such that for each $i < \alpha$ it holds that $C_i[\Box] = Z(t_1, \ldots, t_m)$ with $t_j = \Box$ for exactly one $1 \leq j \leq m$.

The meta-term s is said to satisfy the *finite chains property* if no infinite chain of meta-variables occurs in s.

Example 3.25. An example of a class of meta-terms satisfying the finite chains property is the class of finite meta-terms. The class of meta-terms with infinitely nested chains of finite chains of meta-variables 'guarded' by abstractions or function symbols also satisfies the finite chains property. The following meta-term is an example of a meta-term in the latter class:

$$[x_1]Z_1([x_2]Z_2(\ldots [x_n]Z_n(\ldots)))$$

As a special case we have that any meta-term in which all meta-variables occur as $Z(s_1, \ldots, s_n)$ with no meta-variables occurring at the roots of $s_1, \ldots, s_m$ satisfies the finite chains property.

Examples of meta-terms that do *not* satisfy the finite chains property are $Z(Z(\ldots Z(\ldots)))$ and $Z_1(Z_2(\ldots Z_n(\ldots)))$.

For later use, we have the following for meta-terms satisfying the finite chains property.

Proposition 3.26. *Let s be a meta-term satisfying the finite chains property and let γ be a map that assigns to each $p \in \mathcal{P}os(s)$ the number of prefix positions of p at which* no *meta-variable occurs. For any $n \in \mathbb{N}$, the number of positions p with $\gamma(p) = n$ is finite.*

Proof. Consider s as a finitely-branching tree. Remove from this tree all positions p for which $\gamma(p) > n$. As any suffix p' of such a p must also satisfy $\gamma(p') > n$, the graph resulting from this removal is again a tree, denote it by $\mathcal{T}$.

Assume that $\mathcal{T}$ contains an infinite path, such that for every position p along the path $\gamma(p) \leq n$. Since non-meta-variables can only occur at n positions along the path, there exists a position q such that only meta-variables occur at suffixes of q, contradicting the finite chains property. Hence, no infinite path occurs in $\mathcal{T}$. As $\mathcal{T}$ is finitely branching, König's Lemma yields that $\mathcal{T}$ is finite, implying that the number of positions p for which $\gamma(p) \leq n$, and *a fortiori* $\gamma(p) = n$ is also finite. □

We next show that all valuations are total on the set of meta-terms satisfying the finite chains property.

Proposition 3.27. *Let s be a meta-term satisfying the finite chains property and let $\bar{\sigma}$ a valuation. There exists a unique term that is the result of applying $\bar{\sigma}$ to s.*

Proof. Viewing the application of $\bar{\sigma}$ in the two-step operational way described in the previous section, it is immediate by the definition of valuations that the first step of applying $\bar{\sigma}$ to s has a unique result. Denote this result by s_σ and denote the set containing all parallel β-redexes in s_σ by $\mathcal{U}$. The result now follows if we can show that $\mathcal{U}$ has a complete development ending in a term and, moreover, that each development of $\mathcal{U}$ ends in the same term.

Although omitted, the definitions of (complete) developments can be easily derived from the definitions for iλc or those presented in Section 6.

To start, observe that to repeatedly rewrite the root of s_σ by means of the parallel β-redex requires the root to be of the form

$$(\underline{\lambda}\vec{x}.x_i)(t_1, \ldots, t_n)\,,$$

where $1 \leq i \leq n$ and t_i is again such a redex. This is only possible if there exists in s_σ an infinite chain of such redexes starting at the root. However, this means that an infinite chain of meta-variables must occur in s, which is impossible as s satisfies the finite chains condition. Thus, the root can only be rewritten finitely often in a development. Applying the same reasoning to the roots of the subterms, we obtain a complete development reducing the redexes of $\mathcal{U}$ in an outside-in fashion. Since all parallel β-redexes occur in $\mathcal{U}$ and since no $\underline{\lambda}$-applications and $\underline{\lambda}$-abstractions occur in s, the result of the complete development, which we denote $\bar{\sigma}(s)$, is necessarily a term.

To show that each complete development ends in $\bar{\sigma}(s)$, observe that we can consider each parallel β-redex $(\underline{\lambda}x_1, \ldots, x_n.s)(t_1, \ldots, t_n)$ to be a sequence of β-redexes:

$$(\lambda x_1.(\ldots((\lambda x_n.s)t_n)\ldots))t_1\,.$$

This means that each complete development in our variant of iλc corresponds to a complete development in iλc extended with some function symbols. As each complete development in iλc ends in the same term [16, 17], a result independent of any added function symbols, the same holds for the redexes in $\mathcal{U}$. Hence, $\bar{\sigma}(s)$ is unique. □

Remark 3.28. The subset of meta-terms satisfying the finite chains property can alternatively be defined by slightly altering the depth measure and metric employed to define infinite meta-terms.

Given a term s and a position $p \in \mathcal{P}os(s)$, define the depth measure D:

$$\begin{aligned} D(s, \epsilon) &= 0 \\ D(Z(t_1, \ldots, t_n), i \cdot p') &= D(t_i, p') \\ D([x]t, 0 \cdot p') &= 1 + D(t, p') \\ D(f(t_1, \ldots, t_n), i \cdot p') &= 1 + D(t_i, p') \end{aligned}$$

The difference with the usual depth measure $|p|$ is to be found in the fact that we are *not* counting meta-variables towards the depth.

Next define the metric d_a:

$$d_a(s,t) = \begin{cases} 0 & \text{if } s \text{ and } t \text{ are } \alpha\text{-equivalent} \\ 2^{-k} & \text{otherwise} \end{cases}$$

where k is the minimal depth — with respect to the depth measure D — such that a conflict occurs between s and t.

The meta-terms without infinite chains of meta-variables are now defined by taking the metric completion of the set of finite meta-terms with respect to d_a. The conclusion that precisely the meta-terms without infinite chains of meta-variables are obtained is an immediate consequence of the meta-variables not counting towards the depth.

The above construction for the subset of meta-terms satisfying the finite chains property is inspired by similar constructions for iλc defining subsets of the set of infinite λ-terms by slightly altering the notion of the depth measure used in the employed metric [16]. The set containing no λ-terms with infinite chains of λ-abstractions (i.e. subterms of the form $\lambda x_1.\lambda x_2 \ldots \lambda x_n \ldots$) can e.g. be defined in this way.

4 Rewrite Rules and Reductions

Having defined terms, we move on to define the rewrite relation and reductions of terms.

4.1 Rewrite Rules

We give a number of definitions that are direct extensions of the corresponding definitions from CRS theory.

Definition 4.1. A finite meta-term is a *pattern* if each of its meta-variables has distinct bound variables as its arguments. Moreover, a meta-term is *closed* if all its variables occur bound.

We next define rewrite rules and iCRSs. As in the case of iTRSs, the definitions are identical to the definitions in the finite case, with exception of the restrictions on the right-hand sides of the rewrite rules [9, 18]. In the case of iTRSs the finiteness restriction is lifted from the right-hand sides. Here, this is also done, but at the same time the finite chains property is put in place.

Definition 4.2. A *rewrite rule* is a pair (l, r), denoted $l \to r$, where l is a finite meta-term and r is a meta-term, such that:

1. l is a pattern and of the form $f(s_1, \ldots, s_n)$ with $f \in \Sigma$ of arity n,
2. all meta-variables that occur in r also occur in l,
3. l and r are closed, and
4. r satisfies the finite chains property.

The meta-terms l and r are called, respectively, the *left-hand side* and the *right-hand side* of the rewrite rule.

An *infinitary Combinatory Reduction System (iCRS)* is a pair $\mathcal{C} = (\Sigma, R)$ with Σ a signature and R a set of rewrite rules.

With respect to the left-hand sides of rewrite rules, it is always the case that only finite chains of meta-variables occur, since the left-hand sides are finite. Moreover, it follows easily that iTRSs and iλc are iCRSs if we interpret their rewrite rules as rules in the above sense. By definition of iTRSs and iλc we have that only finite chains of meta-variables occur in the right-hand sides of the rewrite rules.

We now define rewrite steps.

Definition 4.3. A *rewrite step* is a pair of terms (s, t), denoted $s \to t$, such that $s = C[\bar{\sigma}(l)]$ and $t = C[\bar{\sigma}(r)]$ for some one-hole context $C[\Box]$, rewrite rule $l \to r$, and valuation $\bar{\sigma}$. The term $\bar{\sigma}(l)$ is called an $l \to r$-*redex*, or simply a *redex*. The redex *occurs* at position p and depth $|p|$ in s, where p is the position of hole in $C[\Box]$.

Any position q of s is said to occur in the *redex pattern* of the redex at position p if $q \geq p$ and if there does not exist a position q' with $q \geq p \cdot q'$ such that q' is the position of a meta-variable in l.

We now mention some standard restrictions on rewrite rules that we shall need later in the paper:

Definition 4.4. A rewrite rule is *left-linear*, if each meta-variable occurs at most once in its left-hand side. Moreover, an iCRS is *left-linear* if all its rewrite rules are.

Definition 4.5. A pattern is *fully-extended* [14, 37], if, for each of its meta-variables Z and each abstraction $[x]s$ having an occurrence of Z in its scope, x is an argument of that occurrence of Z. Moreover, a rewrite rule is *fully-extended* if its left-hand side is and an iCRS is *fully-extended* if all its rewrite rules are.

Definition 4.6. Let s and t be finite meta-terms that have no meta-variables in common. The meta-term s *overlaps* t if there exists a non-meta-variable position $p \in \mathcal{P}os(s)$ and a valuation $\bar{\sigma}$ such that $\bar{\sigma}(s|_p) = \bar{\sigma}(t)$.

Two rewrite rules overlap if their left-hand sides overlap and if the overlap does not occur at the root when two copies of the same rule are considered. An iCRS is *orthogonal* if all its rewrite rules are left-linear and no pair of rewrite rules overlaps.

Remark 4.7. It is easily seen that if two left-linear rules overlap in an infinite term, there is also a finite term in which they overlap. As left-hand sides are *finite* meta-terms, we may appeal to standard ways of deeming CRSs orthogonal by inspection of their rules. We shall do so informally on several occasions in the remainder of the paper.

Definition 4.8. A rewrite rule is *collapsing* if its right-hand side has a meta-variable at the root. Moreover, a redex and a rewrite step are *collapsing* if the employed rewrite rule is. A redex is called *root-collapsing* if it is collapsing and occurs at position ϵ.

4.2 Transfinite Reductions

We can now define transfinite reductions. The definition is equivalent to those for iTRSs and iλc [18, 16].

Definition 4.9. A *transfinite reduction* with domain $\alpha > 0$ is a sequence of terms $(s_\beta)_{\beta<\alpha}$ such that $s_\beta \to s_{\beta+1}$ for all $\beta + 1 < \alpha$. In case $\alpha = \alpha' + 1$, the reduction is *closed* and of length α'. In case α is a limit ordinal, the reduction is called *open* and of length α. The reduction is *weakly continuous* or *Cauchy continuous* if, for limit every ordinal $\gamma < \alpha$, the distance between s_β and s_γ tends to 0 as β approaches γ from below. The reduction is *weakly convergent* or *Cauchy convergent* if it is weakly continuous and closed.

Intuitively, an open transfinite reduction is lacking a well-defined final term, while a closed reduction does have such a term.

As in [18, 16, 17], we prefer to reason about strongly convergent reductions. This ensures that we can restrict our attention to reductions of length at most ω by the so-called *compression property*, as shown in Section 5.

Definition 4.10. Let $(s_\beta)_{\beta<\alpha}$ be a transfinite reduction. For each rewrite step $s_\beta \to s_{\beta+1}$, let d_β denote the depth of the contracted redex. The reduction is *strongly continuous* if it is weakly continuous and if, for every limit ordinal $\gamma < \alpha$, the depth d_β tends to infinity as β approaches γ from below. The reduction is *strongly convergent* if strongly continuous and closed.

Notation 4.11. By $s \twoheadrightarrow^{\alpha} t$, respectively $s \twoheadrightarrow^{\leq\alpha} t$, we denote a *strongly convergent* reduction of ordinal length α, respectively of ordinal length less than or equal to α. By $s \twoheadrightarrow t$ we denote a *strongly convergent* reduction of arbitrary ordinal length and by $s \to^{*} t$ we denote a reduction of finite length. Reductions are usually ranged over by capitals such as D, S, and T. The concatenation of reductions S and T is denoted by $S;T$.

Note that the concatenation of any finite number of strongly convergent reductions is a strongly convergent reduction. With respect to strongly convergent reductions we also have the following:

Lemma 4.12. *If $s \twoheadrightarrow t$, then the number of steps contracting redexes at depths less than $d \in \mathbb{N}$ is finite for any d.*

Proof. This is exactly the proof of [18, Lemma 3.5]. □

Corollary 4.13. *Every strongly convergent reduction has countable length.*

4.3 Descendants and Residuals

The definition of a descendant across a rewrite step $\bar{\sigma}(l) \to \bar{\sigma}(r)$ follows the definition of substitution, and is thus defined in two steps. The first step defines descendants in $\bar{\sigma}(r)$ where only the valuation is applied and not Equation (1). The second step defines descendants across application of Equation (1).

Given that the second step of the substitution is just a complete development in a variant of iλc, the second step in the definition of descendants is just a variant of descendants in iλc [16, 17]. For this reason, the second step is not made explicit here. However, it should be remarked that we use a variant of descendants in iλc in which positions of variables bound by parallel β-redexes that are being reduced do not have any descendants, while the behaviour with respect to all other positions is a usual. As a consequence of this assumption, positions of variables bound by redexes being reduced in iCRSs will not have descendants either. This behaviour is analogous to that of descendants defined in [22].

We next give a definition for the first step. In the definition we denote by 0 the position of the subterm on the left-hand side of a $\underline{\lambda}$-application and also the position of the body of a $\underline{\lambda}$-abstraction. By $1, \ldots, n$ we denote the positions of the subterms on the right-hand side of the $\underline{\lambda}$-application. This means that $(\underline{\lambda}\vec{x}.s)(t_1, \ldots, t_n)|_0 = (\underline{\lambda}\vec{x}.s)$, $\underline{\lambda}\vec{x}.s|_0 = s$, and $Z(t_1, \ldots, t_n)|_i = (\underline{\lambda}\vec{x}.s)(t_1, \ldots, t_n)|_i = t_i$ for $1 \leq i \leq n$. We denote by $\bar{\sigma}(l) \to r_\sigma$ the rewrite step $\bar{\sigma}(l) \to \bar{\sigma}(r)$ where the valuation is applied to r but not Equation (1).

Definition 4.14. Let $l \to r$ be a rewrite rule, $\bar{\sigma}$ a valuation, and $p \in \mathcal{P}os(\bar{\sigma}(l))$. Suppose $u : \bar{\sigma}(l) \to r_\sigma$. The set $p/_1 u$ is defined as follows:

- if a position $q \in \mathcal{P}os(l)$ exists such that $p = q \cdot q'$ and $root(l|_q) = Z$, then define $p/_1 u = \{p' \cdot 0 \cdot 0 \cdot q' \mid p' \in P\}$ with $P = \{p' \mid root(r|_{p'}) = Z\}$,
- if no such position exists, then define $p/_1 u = \emptyset$.

Note that $\mathcal{P}os(r) \subseteq \mathcal{P}os(r_\sigma)$ by the notation of positions in subterms of the form $(\underline{\lambda}\vec{x}.s)(t_1, \ldots, t_n)$. From this it follows that $P \subseteq \mathcal{P}os(r_\sigma)$.

We can now give the full definition of a descendant across a rewrite step.

Definition 4.15. Let $u : C[\bar{\sigma}(l)] \to C[\bar{\sigma}(r)]$ be a rewrite step, such that p is the position of the hole in $C[\Box]$, and let $q \in \mathcal{P}os(C[\bar{\sigma}(l)])$. The set of *descendants* of q *across* u, denoted q/u, is defined as $q/u = \{q\}$ in case $p \parallel q$ or $p < q$. In case $q = p \cdot q'$, it is defined as $q/u = \{p \cdot q'' \mid p'' \in Q\}$, where Q is the set of descendants of $q'/_1 u'$ with $u' : \bar{\sigma}(l) \to r_\sigma$ across a complete development of the parallel β-redexes in r_σ.

Descendants across a reduction are defined as for iTRSs and iλc.

Definition 4.16. Let $s_0 \twoheadrightarrow^\alpha s_\alpha$ and let $P \subseteq \mathcal{P}os(s_0)$. The set of *descendants* of P across $s_0 \twoheadrightarrow^\alpha s_\alpha$, denoted $P/(s_0 \twoheadrightarrow^\alpha s_\alpha)$, is defined as follows:

- if $\alpha = 0$, then $P/(s_0 \twoheadrightarrow^\alpha s_\alpha) = P$,
- if $\alpha = 1$, then $P/(s_0 \to s_1) = \bigcup_{p \in P} p/(s_0 \to s_1)$,
- if $\alpha = \beta + 1$, then $P/(s_0 \twoheadrightarrow^{\beta+1} s_{\beta+1}) = (P/(s_0 \twoheadrightarrow^\beta s_\beta))/(s_\beta \to s_{\beta+1})$,
- if α is a limit ordinal, then $p \in P/(s_0 \twoheadrightarrow^\alpha s_\alpha)$ iff $p \in P/(s_0 \twoheadrightarrow^\beta s_\beta)$ for all large enough $\beta < \alpha$.

In the case of orthogonal iCRSs, if there exists a redex at a position p employing a rewrite rule $l \to r$ that is not contracted in a rewrite step and if p has descendants across the step, then there exists a redex at each descendant of p that also employs the rule $l \to r$. Hence, for orthogonal systems there exists a well-defined notion of *residual* by strongly convergent reductions. We overload the notation $\cdot/\cdot$ to denote both the descendant and the residual relation.

Notation 4.17. Let $s \twoheadrightarrow t$. Assume $P \subseteq \mathcal{P}os(s)$ and $\mathcal{U}$ a set of redexes in s. We denote the descendants of P across $s \twoheadrightarrow t$ by $P/(s \twoheadrightarrow t)$ and the residuals of $\mathcal{U}$ across $s \twoheadrightarrow t$ by $\mathcal{U}/(s \twoheadrightarrow t)$. Moreover, if $P = \{p\}$ and $\mathcal{U} = \{u\}$, then we also write $p/(s \twoheadrightarrow t)$ and $u/(s \twoheadrightarrow t)$. Finally, if $s \twoheadrightarrow t$ consists of a single step contracting a redex u, then we sometimes write $\mathcal{U}/u$.

$$
\begin{array}{ccc}
\sigma_0(l) & \xrightarrow{\leq\omega}\!\!\!\!\twoheadrightarrow & \sigma_\omega(l) \\
\downarrow & & \downarrow \\
\sigma_0(r) & \xrightarrow{\leq\omega}\!\!\!\!\twoheadrightarrow & \sigma_\omega(r)
\end{array}
$$

Figure 4: Lemma 5.1

The following two lemmas provide some insight in the interplay between residuals and strongly convergent reductions, they are the respective analogues of Lemmas 12.5.12 and 12.5.4 in [17].

Lemma 4.18. *Let $\mathcal{U}$ be a set of positions in a term s and let $s \twoheadrightarrow t$. If every step in $s \twoheadrightarrow t$ occurs strictly below depth d, then $\mathcal{U}$ and $\mathcal{U}/(s \twoheadrightarrow t)$ have exactly the same members at depth at most d.*

Proof. No reduction can affect any part of the term at lesser depths than its steps. □

Lemma 4.19. *For every fully-extended, left-linear iCRS, if $s \twoheadrightarrow t$ is a reduction of limit ordinal length, then for every redex u in t there exists a term s' in $s \twoheadrightarrow t$ such that u is the unique residual of a redex in s'.*

Proof. Suppose that u is a redex in t that occurs at position p. By definition of rewrite rules, it follows that the left-hand side of the rewrite rule employed in u is finite. Hence, there exists a depth d such that all positions in the redex pattern of u have depth strictly less than d. By strong convergence we may write $s \twoheadrightarrow t$ as $s \twoheadrightarrow s' \twoheadrightarrow t$ such that all steps in $s' \twoheadrightarrow t$ occur below depth d. By left-linearity and fully-extendedness it now follows that a redex v occurs at position p in s' with u the unique residual of v. □

5 Compression

Compression is a feature of (strongly convergent) infinitary rewriting that, in essence, allows us to 'compress' reductions of arbitrary lengths to much more manageable 'equivalent' reductions of length at most ω.

In this section, we prove the compression property for fully-extended, left-linear iCRSs. Fully-extendedness and left-linearity ensure that no redex is created by either making two subterms equal in an infinite number of steps or by erasing some variable in an infinite number of steps. We will show later in the section that these two assumptions cannot be omitted.

We first prove an auxiliary lemma:

Lemma 5.1. *For every fully-extended, left-linear iCRS, if $\sigma_0(l) \twoheadrightarrow^{\leq\omega} \sigma_\omega(l) \to \sigma_\omega(r)$, then $\sigma_0(l) \to \sigma_0(r) \twoheadrightarrow^{\leq\omega} \sigma_\omega(r)$ (see Figure 4).*

Proof. Let $\sigma_0(l) \twoheadrightarrow^{\leq\omega} \sigma_\omega(l) \to \sigma_\omega(r)$. By left-linearity and fully-extendedness we have that $\sigma_0(l) \to \sigma_0(r)$. Hence, $\sigma_0(r) \twoheadrightarrow^{\leq\omega} \sigma_\omega(r)$ is left to prove.

Since the left-hand side of each rewrite rule is a pattern, it follows that $\sigma_0(l) \twoheadrightarrow \sigma_\omega(l)$ consists of a finite number of interleaved, strongly convergent reductions of length at most ω: one reduction for each meta-variable Z that occurs in l reducing $\sigma_0(Z)(\vec{x})$ to $\sigma_\omega(Z)(\vec{x})$. By Lemma 4.12 we may write:

$$\sigma_0(Z)(\vec{x}) \to^* \sigma_1(Z)(\vec{x}) \to^* \cdots \to^* \sigma_d(Z)(\vec{x}) \to^* \sigma_{d+1}(Z)(\vec{x}) \to^* \cdots \sigma_\omega(Z)(\vec{x}),$$

where for each $d \geq 0$ we have that all steps in $\sigma_d(Z)(\vec{x}) \twoheadrightarrow \sigma_\omega(Z)(\vec{x})$ occur at depth d or below. Hence, $\sigma_d(Z)(\vec{x}) \twoheadrightarrow \sigma_\omega(Z)(\vec{x})$ is possibly empty. Moreover, by left-linearity, we may replace the variables $\vec{x}$ by arbitrary terms $\vec{t}$ to obtain a reduction $\sigma_d(Z)(\vec{t}) \to^* \sigma_{d+1}(Z)(\vec{t})$. No nesting of the terms in $\vec{t}$ can occur, as the free variables in $\vec{t}$ are also free in $\sigma_d(Z)(\vec{t})$.

We now show for all $d \geq 0$ that there exists a reduction $s_d \to^* s_{d+1}$ with all rewrite steps occurring at depth d or below and such that $d(s_d, \sigma_\omega(r)) \leq 2^{-d}$ and $s_0 = \sigma_0(r)$. To do so, define a map γ that assigns to each $p \in \mathcal{P}os(r)$ the number of prefix positions of p at which *no* meta-variable occurs. By Proposition 3.26 we have for any $n \in \mathbb{N}$ that the number of positions p with $\gamma(p) = n$ is finite.

Label each meta-variable in r with its position yielding a labelled meta-term r'. Denote the labelled meta-variables in r' by Z^p and define the following for each $d \geq 0$ and Z^p:

$$\sigma'_d(Z^p) = \begin{cases} \sigma_0(Z) & \text{if } d \leq \gamma(p) \\ \sigma_{d-\gamma(p)}(Z) & \text{if } d > \gamma(p) \end{cases}$$
$$s_d = \sigma'_d(r')$$

with the final σ'_d the valuation induced by the map defined on meta-variables.

For each Z^p with $d > \gamma(p)$ consider $\sigma_{d-\gamma(p)}(Z)(\vec{x}) \to^* \sigma_{d-\gamma(p)+1}(Z)(\vec{x})$. As the number of meta-variable positions with $d > \gamma(p)$ is finite and as no new nestings can be created, it follows that $s_d \to^* s_{d+1}$. Since all steps in $\sigma_{d-\gamma(p)}(Z)(\vec{x}) \to^* \sigma_{d-\gamma(p)+1}(Z)(\vec{x})$ for Z at position p in r occur at depth $d - \gamma(p)$ or below and since there are $\gamma(p)$ prefix positions of p at which no meta-variable occurs, all rewrite steps in $s_d \to^* s_{d+1}$ occur at depth d or below. Moreover, since all rewrite steps in $\sigma_{d-\gamma(p)}(Z)(\vec{x}) \twoheadrightarrow \sigma_\omega(Z)(\vec{x})$ also occur at depth $d - \gamma(p)$ or below, we also have $d(s_d, \sigma_\omega(r)) \leq 2^{-d}$.

By construction of $s_d \to^* s_{d+1}$ it follows that

$$s_0 \to^* s_1 \to^* \cdots \to^* s_d \to^* s_{d+1} \to^* \cdots \sigma_\omega(r)$$

is a strongly convergent reduction of length at most ω. Moreover, as $\sigma_0(r) = s_0$, we have $\sigma_0(r) \twoheadrightarrow^{\leq\omega} \sigma_\omega(r)$. Hence, the result now follows. □

The main result of the section is now as follows:

Theorem 5.2 (Compression)**.** *For every fully-extended, left-linear iCRS, if* $s \twoheadrightarrow^{\alpha} t$*, then* $s \twoheadrightarrow^{\leq\omega} t$*.*

Proof. Let $s \twoheadrightarrow^{\alpha} t$ and proceed by ordinal induction on α. By [17, Theorem 12.7.1] it suffices to show that the theorem holds for $\alpha = \omega + 1$: The cases where α is 0, a limit ordinal, or a successor ordinal greater than $\omega + 1$ do not depend on the definition of rewriting.

Suppose $\alpha = \omega + 1$ and write

$$s = s_0 \to s_1 \to \cdots s_\omega \to s_{\omega+1} = t\,.$$

The redex contracted in $s_\omega \to s_{\omega+1}$, call it u, occurs at a position p at depth d_u in s_ω. By definition of rewrite rules, the rule employed in u, say $l \to r$, has a finite left-hand side. Hence, there exists a $d_l > d_u$ such that all positions in the redex pattern of u have depth strictly less than d_l.

By Lemma 4.12, we may write $s \twoheadrightarrow t$ as:

$$s_0 \to^* s_n \twoheadrightarrow s_\omega \to s_{\omega+1}$$

where all rewrite steps in $s_n \twoheadrightarrow s_\omega$ occur at depth d_l or below. Moreover, by left-linearity and fully-extendedness it follows that a redex v occurs at position p in s_n with u the unique residual of v. Contracting v in s_n yields a term t'.

Observe for some $m \in \mathbb{N}$ there exists a context $C[\Box, \ldots, \Box]$ with $m+1$ holes, which all occur at depth d_u, such that we may write $s_n \twoheadrightarrow s_\omega \to s_{\omega+1}$ as:

$$C[\sigma(l), s'_1, \ldots, s'_m] \twoheadrightarrow C[\sigma'(l), s''_1, \ldots, s''_m] \to C[\sigma'(r), s''_1, \ldots, s''_m]\,.$$

Existence follows as each rewrite step in $s_n \twoheadrightarrow s_\omega$ occurs at depth $d_l > d_u$ or below and as all positions in the redex pattern of redex v occur at or at depth d_u or below.

By definition of $C[\Box, \ldots, \Box]$ we have that $t' = C[\sigma(r), s'_1, \ldots, s'_m]$, where t' is the result of contracting v in s_n. Moreover, the reduction $s_n \twoheadrightarrow s_{\omega+1}$ interleaves the reductions $\sigma(l) \twoheadrightarrow^{\leq\omega} \sigma'(l) \to \sigma'(r)$ and $s'_i \twoheadrightarrow^{\leq\omega} s''_i$, with $1 \leq i \leq m$, where for the first of these reductions, there exists a reduction $\sigma(l) \to \sigma(r) \twoheadrightarrow^{\leq\omega} \sigma'(r)$ by Lemma 5.1.

By Lemma 4.12, we may write $\sigma(r) \twoheadrightarrow^{\leq\omega} \sigma'(r)$ as:

$$\sigma(r) = \sigma^0(r) \to^* \sigma^1(r) \to^* \cdots \to^* \sigma^d(r) \to^* \sigma^{d+1}(r) \to^* \cdots \sigma'(r)$$

and each $s'_i \twoheadrightarrow^{\leq\omega} s''_i$ as:

$$s'_i = s^0_i \to^* s^1_i \to^* \cdots \to^* s^d_i \to^* s^{d+1}_i \to^* \cdots s''_i\,,$$

where for each $d \geq 0$ we have that steps in $\sigma^d(r) \twoheadrightarrow \sigma'(r)$ and $s^d_i \twoheadrightarrow s''_i$ occur at depth d or below. Hence, $\sigma^d(r) \twoheadrightarrow \sigma'(r)$ and $s^d_i \twoheadrightarrow s''_i$ may be empty from some d onwards.

We now show for all $d \geq 0$ that there exists a reduction $t_d \to^* t_{d+1}$ with all rewrite steps occurring at depth $d_u + d$ or below and such that $d(t_d, t) \leq 2^{-(d_u+d)}$. To do so, define the following for each $d \geq 0$:

$$t_d = C[\sigma^d(r), s^d_1, \ldots, s^d_m]$$

and consider $\sigma^d(r) \to^* \sigma^{d+1}(r)$ and $s_i^d \to^* s_i^{d+1}$ for all $1 \leq i \leq m$. Obviously, we have:

$$t_d = C[\sigma^d(r), s_1^d, \ldots, s_m^d] \to^* C[\sigma^{d+1}(r), s_1^{d+1}, \ldots, s_m^{d+1}] = t_{d+1}\,.$$

Since all steps in $\sigma^d(r) \to^* \sigma^{d+1}(r)$ and $s_i^d \to^* s_i^{d+1}$ occur at depth d or below and since the holes in the context $C[\square, \ldots, \square]$ occur at depth d_u, all rewrite steps in $t_d \to t_{d+1}$ occur at depth $d_u + d$ or below. Moreover, since all rewrite steps in $\sigma^d(r) \twoheadrightarrow \sigma'(r)$ and $s_i^d \twoheadrightarrow s_i''$ also occur at depth d or below, we also have that $d(t_d, t) \leq 2^{-(d_u+d)}$.

By construction of the reductions $t_d \to^* t_{d+1}$ it follows that

$$t_0 \to^* t_1 \to^* \cdots \to^* t_d \to^* t_{d+1} \to^* \cdots t$$

is a strongly convergent reduction of length at most ω. Since $t' = t_0$, we have that $t' \twoheadrightarrow^{\leq\omega} t$. Hence, as $s \to^* t'$, it follows that $s \twoheadrightarrow^{\leq\omega} t$, as required. $\square$

The previous proof is completely independent of the particulars of the notion of rewriting involved, as long as it is based on terms and contexts. Indeed, the proof is essentially a spelled out and detailed version of earlier proofs of compression properties in more restricted settings, e.g. iλc [16]. The details specific to iCRSs are restricted to Lemma 5.1.

We next show that the assumptions of left-linearity and fully-extendedness cannot be omitted from the previous theorem. In addition we show that omitting the finite chains property from the definition of a rewrite rule can also make compression fail.

Example 5.3 (Failure of compression without left-linearity). In case left-linearity is omitted, failure of compression follows if we interpret the counterexample to compression for non-left-linear iTRSs [18] in the context of iCRSs. That is, suppose we have at our disposal the following three rewrite rules:

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

Obviously, the first of the above rules is not left-linear. Now consider the following reduction of length $\omega + 1$:

$$f(a, b) \to^* f(g(a), g(b)) \to^* f(g^2(a), g^2(b)) \to^* \cdots f(g^\omega, g^\omega) \to c$$

The reduction cannot the compressed to a reduction of length at most ω, because ω steps are required to reduce both $g(a)$ and $g(b)$ to g^ω and because the two arguments of f differ as long as $g(a)$ and $g(b)$ have not been reduced to g^ω.

Example 5.4 (Failure of compression without fully-extendedness). Consider the following two rewrite rules:

$$\begin{aligned} f([x]Z) &\to Z \\ g(Z) &\to h(g(Z)) \end{aligned}$$

The first of the above two rewrite rules is not fully-extended, as the meta-variable Z on the left-hand side occurs in the scope of the abstraction $[x]$, while x is not an argument of Z. Now consider the following reduction:

$$f([x]g(x)) \to f([x]h(g(x))) \to \cdots f([x]h^\omega) \to h^\omega$$

The reduction cannot be compressed to a reduction of length at most ω, because ω steps are required to reduce $g(x)$ to h^ω and because the variable x occurs bound as long as $g(x)$ has not been reduced to h^ω.

Alternative to the above, failure of compression in the case of non-fully-extendedness also follows by interpreting the $\lambda\beta\eta$-calculus in the context of iCRSs. The η-rule is not fully-extended. That compression to reductions of at most length ω fails is demonstrated in [16]. However, as shown in [33] a slightly different compression property does hold in the case of the $\lambda\beta\eta$-calculus: Each reduction can be compressed to a reduction of length at most $\omega + \omega$.

Lastly, we show that the finite chains property that underlies much of treatment of iCRSs is also needed for compression:

Example 5.5 (Failure of compression without the finite chains property). Assume we have at our disposal the the following two rewrite rules:

$$\begin{aligned} f([x]Z(x), [y]Z'(y)) &\to Z'(Z^\omega) \\ g(Z) &\to h(g(Z)) \end{aligned}$$

Obviously, the right-hand side of the first rule does not satisfy the finite chains property. Now consider the following reduction:

$$f([x]x, [y]g(y)) \to f([x]x, [y]h(g(y))) \to \cdots f([x]x, [y]h^\omega) \to h^\omega$$

Compression fails, as the first rule cannot be applied to $f([x]x, [y]g(y))$, or for that matter to any $f([x]x, [y]h^n(g(y)))$ with $n \in \mathbb{N}$, because we have:

$$f([x]x, [y]h^n(g(y))) = \bar{\sigma}(f([x]Z(x), [y]Z'(y)))\,,$$

with $\sigma(Z) = \underline{\lambda}x.x$ and $\sigma(Z') = \underline{\lambda}y.h^n(g(y))$, and:

$$\bar{\sigma}(Z'(Z^\omega)) = (\underline{\lambda}y.h^n(g(y)))((\underline{\lambda}x.x)((\underline{\lambda}x.x)(\ldots((\underline{\lambda}x.x)(\ldots)))))\,,$$

which obviously has no complete development of its parallel β-redexes.

6 Developments

In this section we prove that each complete development of the same set of redexes in an orthogonal iCRS ends in the same term.

Assuming in the remainder of this section that every iCRS is orthogonal and that s is a term and $\mathcal{U}$ a set of redexes in s, we first define developments:

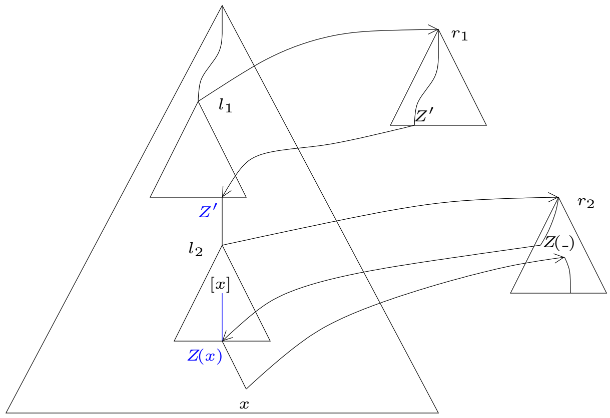


Figure 5: A path tracing trough a term: when a redex or a bound variable is met in a trace, a 'jump' is made to the right-hand side of the employed rewrite rule and the trace continues there until a meta-variable is encountered.

Definition 6.1. A *development* of $\mathcal{U}$ is a strongly convergent reduction such that each step contracts a residual of a redex in $\mathcal{U}$. A development $s \twoheadrightarrow t$ is called *complete* if $\mathcal{U}/(s \twoheadrightarrow t) = \emptyset$. Moreover, a development is called *finite* if $s \twoheadrightarrow t$ is finite.

Remark 6.2. Although the above the definition and the results below concern orthogonal iCRSs, they can actually be interpreted in the more liberal context of orthogonal sets of redexes, i.e. where no restrictions are placed upon the iCRSs but where it is assumed that there is no overlap between the redexes that occur in $\mathcal{U}$. No modification of either the above definition or the proofs below is necessary. Even though this is the case, we opt to work in the context of orthogonal iCRSs, as this is most common throughout the literature on rewrite systems.

6.1 Paths and Finite Jumps

To prove that each complete development of the same set of redexes ends in the same term, we extend the technique of the Finite Jumps Developments Theorem [17] to orthogonal iCRSs. The theorem employs the notions of paths and path projections. In essence, paths and path projections 'trace' through terms starting at the root and proceeding to increasingly greater depths. Most importantly, if a redex to be contracted in a development or a variable bound by such a redex is met in a trace, a 'jump' is made to the right-hand side of the employed rewrite rule. The trace continues there until a meta-variable is met, at which point a jump back to the original term is made (see Figure 5).

In the following, we denote by p_u the position of the redex u in s. Moreover, we say that a variable x is *bound by a redex* u if x is bound by an abstraction $[x]$ which occurs in the left-hand side of the rewrite rule employed in u.

Definition 6.3. A *path* of s with respect to $\mathcal{U}$ is a sequence of nodes and edges. Each node is labelled either (s, p) with $p \in \mathcal{P}os(s)$ or (r, p, p_u) with r the right-hand side of a rewrite rule, $p \in \mathcal{P}os(r)$, and $u \in \mathcal{U}$. Each directed edge is either unlabelled or labelled with an element of $\mathbb{N}$.

Every path starts with a node labelled (s, ϵ). If a node n of a path is labelled (s, p) and has an outgoing edge to a node n', then:

1. if the subterm at p is not a redex in $\mathcal{U}$, then for some $i \in \mathcal{P}os(s|_p) \cap \mathbb{N}$ the node n' is labelled $(s, p \cdot i)$ and the edge from n to n' is labelled i,
2. if the subterm at p is a redex $u \in \mathcal{U}$ with $l \to r$ the employed rewrite rule, then the node n' is labelled (r, ϵ, p_u) and the edge from n to n' is unlabelled,
3. if $s|_p$ is a variable x bound by a redex $u \in \mathcal{U}$ with $l \to r$ the employed rewrite rule, then the node n' is labelled $(r, p' \cdot i, p_u)$ and the edge from n to n' is unlabelled, such that (r, p', p_u) was the last node before n with p_u, $root(r|_{p'}) = Z$, the unique position of Z in l is q, and $l|_{q \cdot i} = x$.

If a node n of a path is labelled (r, p, p_u) and has an outgoing edge to a node n', then:

1. if $root(r|_p)$ is not a meta-variable, then for some $i \in \mathcal{P}os(r|_p) \cap \mathbb{N}$ the node n' is labelled $(r, p \cdot i, p_u)$ and the edge from n to n' is labelled i,
2. if $root(r|_p)$ is a meta-variable Z, then the node n' is labelled $(s, p_u \cdot q)$ and the edge from n to n' is unlabelled, such that $l \to r$ is the rewrite rule employed in u and such that q is the unique position of Z in l.

We say that a path is *maximal* if it is not a proper prefix of another path. We write a path Π as a (possibly infinite) sequence of alternating nodes and edges $\Pi = n_1 e_1 n_2 \ldots$.

Definition 6.4. Let $\Pi = n_1 e_1 n_2 \ldots$ be a path of s with respect to $\mathcal{U}$. The *path projection* of Π is a sequence of alternating nodes and edges $\phi(\Pi) = \phi(n_1)\phi(e_1)\phi(n_2)\ldots$ such that for each node n of Π:

1. if n is labelled (s, p), then $\phi(n)$ is unlabelled if $s|_p$ is a redex in $\mathcal{U}$ or a variable bound by such a redex and it is labelled $root(s|_p)$ otherwise,
2. if n is labelled (r, p, q), then $\phi(n)$ is unlabelled if $root(r|_p)$ is a meta-variable and it is labelled $root(r|_p)$ otherwise.

For each edge e, if e is labelled i, then $\phi(e)$ has the same label, and if e is unlabelled, then $\phi(e)$ is labelled ϵ.

Example 6.5. Consider the orthogonal iCRS that only has the following rewrite rule, also denoted $l \to r$:

$$f([x]Z(x), Z') \to Z(g(Z(Z'))) .$$

Given the terms $s = f([x]g(x), a)$ and $t = g(g(g(a)))$ and the set $\mathcal{U}$ containing the only redex in s, we have that $s \to t$ is a complete development of $\mathcal{U}$.

The term s has one maximal path with respect to $\mathcal{U}$:

$$(s, \epsilon) \to (r, \epsilon, \epsilon) \to (s, 10) \xrightarrow{1} (s, 101) \to (r, 1, \epsilon) \xrightarrow{1} (r, 11, \epsilon)$$
$$\to (s, 10) \xrightarrow{1} (s, 101) \to (r, 111, \epsilon) \to (s, 2)$$

Moreover, the term t has one maximal path with respect to $\mathcal{U}/(s \to t) = \emptyset$:

$$(t, \epsilon) \xrightarrow{1} (t, 1) \xrightarrow{1} (t, 11) \xrightarrow{1} (t, 111)\,.$$

The path projections of the maximal paths are, respectively,

$$\cdot \xrightarrow{\epsilon} \cdot \xrightarrow{\epsilon} g \xrightarrow{1} \cdot \xrightarrow{\epsilon} g \xrightarrow{1} \cdot \xrightarrow{\epsilon} g \xrightarrow{1} \cdot \xrightarrow{\epsilon} \cdot \xrightarrow{\epsilon} a$$

and

$$g \xrightarrow{1} g \xrightarrow{1} g \xrightarrow{1} a\,.$$

Let $\mathcal{P}(s,\mathcal{U})$ denote the set of path projections of *maximal paths* of s with respect to $\mathcal{U}$. The following two results can be witnessed in the above example, their proofs are simple, but tedious and lengthy, hence occur in Appendix A.

Proposition 6.6. *The map ϕ defines a bijection between the set of paths and the set of path projections, respectively between maximal paths and the path projections in $\mathcal{P}(s,\mathcal{U})$.*

Lemma 6.7. *Let $u \in \mathcal{U}$ and let $s \to t$ be the rewrite step contracting u. There exists a bijection between $\mathcal{P}(s,\mathcal{U})$ and $\mathcal{P}(t,\mathcal{U}/u)$. Given a path projection $\phi(\Pi) \in \mathcal{P}(s,\mathcal{U})$, its image under the bijection is acquired from $\phi(\Pi)$ by deleting finite sequences of unlabelled nodes and ϵ-labelled edges from $\phi(\Pi)$.*

We next define a property $\mathcal{U}$, based on $\mathcal{P}(s,\mathcal{U})$: the *finite jumps property*. We also define some terminology to relate a term to $\mathcal{P}(s,\mathcal{U})$.

Definition 6.8. The set $\mathcal{U}$ has the *finite jumps property* if no path projection occurring in $\mathcal{P}(s,\mathcal{U})$ contains an infinite sequence of unlabelled nodes and ϵ-labelled edges. Moreover, a term t *matches* $\mathcal{P}(s,\mathcal{U})$ if, for all $\phi(\Pi) \in \mathcal{P}(s,\mathcal{U})$ and all prefixes of $\phi(\Pi)$ ending in a node $\phi(n)$ labelled f, it holds that $root(t|_p) = f$, where p is the concatenation of the edge labels in the prefix (starting at the first node of $\phi(\Pi)$ and ending in $\phi(n)$).

We now prove an ancillary result concerning finite jumps; the proof is (almost) identical to the proof of Proposition 12.5.8 in [17].

Proposition 6.9. *If $\mathcal{U}$ has the finite jumps property, then there exists a unique term, denoted $\mathcal{T}(s,\mathcal{U})$, that matches $\mathcal{P}(s,\mathcal{U})$.*

Proof. Let $\mathcal{P}_p(s,\mathcal{U})$ denote the set of all prefixes of path projections in $\mathcal{P}(s,\mathcal{U})$ such that the concatenation of the edge labels for each prefix is p and such that each prefix ends in a labelled node. The proof proceeds by induction on p.

Consider $\mathcal{P}_\epsilon(s,\mathcal{U})$. By the finite jumps property $\mathcal{P}_\epsilon(s,\mathcal{U})$ is non-empty and by the definition of paths, $\mathcal{P}_\epsilon(s,\mathcal{U})$ has at most one element. Hence, $\mathcal{P}_\epsilon(s,\mathcal{U})$ is a singleton set. By definition of paths, the unique prefix in $\mathcal{P}_\epsilon(s,\mathcal{U})$ has precisely one labelled node. Suppose the label is f. It follows that t only matches $\mathcal{P}(s,\mathcal{U})$ if $root(t|_\epsilon) = f$.

Now suppose $\mathcal{P}_p(s,\mathcal{U})$ is a singleton set such that the final node node of the unique prefix in the set is labelled f, where f is either a variable, a function symbol of arity n, or an abstraction. In the last two cases, consider $\mathcal{P}_{p\cdot i}(s,\mathcal{U})$ for either $1 \leq i \leq n$ or $i = 0$. By the finite jumps property, the definition of paths, and the fact that $\mathcal{P}_p(s,\mathcal{U})$ is a singleton set, we have that $\mathcal{P}_{p\cdot i}(s,\mathcal{U})$ is a singleton set. Suppose that final node of the unique prefix in $\mathcal{P}_{p\cdot i}(s,\mathcal{U})$ is labelled g. It follows that t only matches $\mathcal{P}(s,\mathcal{U})$ if $root(t|_{p\cdot i}) = g$.

Since all sets $\mathcal{P}_p(s,\mathcal{U})$ are singleton sets there exist terms that match $\mathcal{P}(s,\mathcal{U})$. Moreover, if t is such a term, then we have for all $p \in \mathcal{P}os(t)$ that $\mathcal{P}_p(s,\mathcal{U})$ exists and is a singleton set (look at all prefixes of p), and if the final labelled node of the unique prefix in such a set has label f, then $root(t|_p) = f$. Hence, the term t is unique. □

We can finally prove the Finite Jumps Developments Theorem:

Theorem 6.10 (Finite Jumps Developments Theorem)**.** *If $\mathcal{U}$ has the finite jumps property, then:*

1. *every complete development of $\mathcal{U}$ ends in $\mathcal{T}(s,\mathcal{U})$,*
2. *for any $p \in \mathcal{P}os(s)$, the set of descendants of p by a complete development of $\mathcal{U}$ is independent of the complete development,*
3. *for any redex u of s, the set of residuals of u by a complete development of $\mathcal{U}$ is independent of the complete development, and*
4. *$\mathcal{U}$ has a complete development.*

Proof. (1) Suppose there is a complete development. We show by ordinal induction that for every s_α in the complete development with residuals $\mathcal{U}_\alpha = \mathcal{U}/(s \twoheadrightarrow s_\alpha)$ of $\mathcal{U}$, we have that $\mathcal{P}(s_\alpha,\mathcal{U}_\alpha)$ can be obtained from $\mathcal{P}(s,\mathcal{U})$ by deleting finite sequences of unlabelled nodes and ϵ-labelled edges from the elements of $\mathcal{P}(s,\mathcal{U})$. Obviously, for $s_0 = s$, this is immediate.

For $s_{\alpha+1}$, it follows by the induction hypothesis that $\mathcal{P}(s_\alpha,\mathcal{U}_\alpha)$ can be obtained from $\mathcal{P}(s,\mathcal{U})$ by deleting finite sequences of unlabelled nodes and ϵ-labelled edges from the elements of $\mathcal{P}(s,\mathcal{U})$. Moreover, by Lemma 6.7, we have that $\mathcal{P}(s_{\alpha+1},\mathcal{U}_{\alpha+1})$ can be obtained from $\mathcal{P}(s_\alpha,\mathcal{U}_\alpha)$ by deleting finite sequences of unlabelled nodes and ϵ-labelled edges. Hence, $\mathcal{P}(s_{\alpha+1},\mathcal{U}_{\alpha+1})$ can also be obtained from $\mathcal{P}(s,\mathcal{U})$ by deleting finite sequences of unlabelled nodes and ϵ-labelled edges.

For s_α with α a limit ordinal, we have by strong convergence that $\mathcal{P}(s_\alpha, \mathcal{U}_\alpha)$ can be obtained from $\mathcal{P}(s, \mathcal{U})$ by deleting all unlabelled nodes and ϵ-labelled edges deleted in the previous steps. As $\mathcal{P}(s, \mathcal{U})$ has the finite jumps property, $\mathcal{P}(s_\alpha, \mathcal{U}_\alpha)$ can only be obtained by deleting finite sequences of unlabelled nodes and ϵ-labelled edges.

Hence, each $\mathcal{P}(s_\alpha, \mathcal{U}_\alpha)$ has the finite jumps property, as $\mathcal{P}(s, \mathcal{U})$ has the finite jumps property and as each $\mathcal{P}(s_\alpha, \mathcal{U}_\alpha)$ can be obtained by deleting only finite sequences of unlabelled nodes and ϵ-labelled edges.

By Proposition 6.9 we have for each $\mathcal{P}(s_\alpha, \mathcal{U}_\alpha)$ that there is a unique term $\mathcal{T}(s_\alpha, \mathcal{U}_\alpha)$ that matches it. By inspection of the proof of the proposition it easily follows that the unlabelled nodes and ϵ-labelled edges are irrelevant for the construction of $\mathcal{T}(s_\alpha, \mathcal{U}_\alpha)$. Hence, $\mathcal{T}(s_\alpha, \mathcal{U}_\alpha) = \mathcal{T}(s, \mathcal{U})$ for all α. Moreover, since the chosen complete development was arbitrary, it follows that the final term of each complete development is $\mathcal{T}(s, \mathcal{U})$.

(2) In analogy to [22, Section II.2], let $\mathcal{K}$ be a set of labels including a special *empty* label ε. Define for all function symbols f, variables x, and for all labels $k \in \mathcal{K}$ the *labelled alternatives* f^k and x^k, where f and f^k have the same arity. A *labelling* of a (meta-)term replaces each function symbol and variable (including the variables that occur in abstractions) by a labelled alternative, assuming that the labels of variables are ignored where *bindings* and *substitutions* are concerned.

The labelled version of the assumed orthogonal iCRS includes for every rewrite rule $l \to r$ and every possible labelling l' of l the rewrite rule $l' \to r'$, where r' is the labelling of r that labels all function symbols and variables with ε. The labelled version of the iCRS is easily shown to be orthogonal (see [22, Proposition II.2.6]).

Each reduction in the labelled version corresponds to a reduction in the original iCRS by removal of all labels. Moreover, given a reduction in the original iCRS and a labelling for the initial term, there exists a unique reduction in the labelled version such that removal of the labels gives the reduction we started out with.

Given a term in which some subterms are labelled k, it is easy to show that the descendants of these subterms across some reduction are precisely the subterms labelled k in the final term. Moreover, these descendants are exactly the descendants obtained in the corresponding unlabelled reduction. The result now follows by the first clause of the current proof when applied to the labelled version of the assumed iCRS, when trivially extended to the slightly more liberal notion of substitutions.

(3) By the second clause of the current proof and orthogonality of the assumed iCRS.

(4) Consider a maximal path Π with a node (s, p) such that p is the position of a redex in $\mathcal{U}$. By definition of paths and path projections we have for the first node $n = (s, p)$ in Π with p the position of a redex that the concatenation of the edge label of the prefix of $\phi(\Pi)$ that ends in $\phi(n)$ is p. Moreover, by inspection of the proof of Lemma 6.7 it follows that contracting the redex at position p deletes a sequence of unlabelled nodes and ϵ-labelled edges from $\phi(\Pi)$ directly

following the node $\phi(n)$.

Repeatedly contract a residual of redex in $\mathcal{U}$ that is at minimal depth. Since only finite sequences of nodes and edges occur and since terms are finitely branching, it follows by the above observations regarding paths that the contracted redexes occur at increasingly greater depths along the constructed reduction. Hence, the reduction is strongly convergent and since redexes at minimal depth are contracted the reduction must also be a complete development. □

Roughly, the above proof is identical to the proof of Proposition 12.5.9 in [17], except that Lemma 6.7 is employed instead of tracing.

6.2 Developments

With the Finite Jumps Developments Theorem in hand, we can now precisely characterise the sets of redexes having complete developments. This characterisation seems to be new.

Recall that we are working with an orthogonal iCRS and that $\mathcal{U}$ is a set of redexes in a term s.

Lemma 6.11. *The set $\mathcal{U}$ has a complete development iff $\mathcal{U}$ has the finite jumps property.*

Proof. To prove that the finite jumps property holds if $\mathcal{U}$ has a complete development, suppose $\mathcal{U}$ does not have the finite jumps property. That is, there exists a path projection $\phi(\Pi)$ of s with respect to $\mathcal{U}$ that ends in an infinite sequence of unlabelled nodes and ϵ-labelled edges.

We show by ordinal induction that for every s_α in the complete development, with residuals $\mathcal{U}_\alpha = \mathcal{U}/(s \twoheadrightarrow s_\alpha)$ of $\mathcal{U}$, that there exists a path projection $\phi(\Pi_\alpha)$ of s_α with respect to $\mathcal{U}_\alpha$ that has an infinite sequence of unlabelled nodes and ϵ-labelled edges. Obviously, for $s_0 = s$, this is immediate.

For $s_{\alpha+1}$, it follows by the induction hypothesis that there exists path projection $\phi(\Pi_\alpha)$ that has an infinite sequence of unlabelled nodes and ϵ-labelled edges. By Lemma 6.7, a path projection $\phi(\Pi_{\alpha+1})$ is obtained by deleting finite sequences of unlabelled nodes and ϵ-labelled edges from $\phi(\Pi)$. Hence, $\phi(\Pi_{\alpha+1})$ must also have an infinite sequence of unlabelled nodes and ϵ-labelled edges.

For s_α with α a limit ordinal, we have by strong convergence that $\phi(\Pi_\alpha)$ can be obtained from $\phi(\Pi)$ by deleting all unlabelled nodes and ϵ-labelled edges deleted in the previous steps. Also by strong convergence, the deleted nodes and edges occur at an increasingly greater distance from the starting node of the path projections considered along the complete development. Hence, an increasing part of the infinite sequence of unlabelled nodes and ϵ-labelled edges is stable along the complete development. This implies that $\phi(\Pi_\alpha)$ also has such an infinite sequence.

Thus, a path projection that has an infinite sequence of unlabelled nodes and ϵ-labelled edges is present after each complete development of $\mathcal{U}$. By definition of paths and path projections this implies that a descendant of a redex in $\mathcal{U}$ occurs in the final term of the complete development. However, this contradicts the fact that no descendants of redexes in $\mathcal{U}$ occur in the final term of a complete development. Hence, $\mathcal{U}$ must have the finite jumps property.

That $\mathcal{U}$ has a complete development if it has the finite jumps property is an immediate consequence of Theorem 6.10(4). □

The result we were aiming at now follows easily.

Theorem 6.12. *If $\mathcal{U}$ has a complete development, then all complete developments of $\mathcal{U}$ end in the same term.*

Proof. By Lemma 6.11, if $\mathcal{U}$ has a complete development, then it has the finite jumps property. But then, each complete development of $\mathcal{U}$ ends in the same term by Theorem 6.10(1). □

6.3 Properties of Developments

We next prove a number of properties of complete developments that will prove to be of use in later sections. Again, recall that we assume we are working in an orthogonal iCRS and that $\mathcal{U}$ is a set of redexes in a term s.

Notation 6.13. If there exists a complete development of $\mathcal{U}$ resulting in a term t, then we write $s \Rightarrow t$, where the arrow is adorned with $\mathcal{U}$ if needed.

Lemma 6.14. *If $\mathcal{U}$ has a complete development and if $s \twoheadrightarrow t$ is a (not necessarily complete) development of $\mathcal{U}$, then $\mathcal{U}/(s \twoheadrightarrow t)$ has a complete development.*

Proof. Since $\mathcal{U}$ has the finite jumps property by Lemma 6.11, it follows by inspection of the proof of Theorem 6.10(1) that $\mathcal{U}/(s \twoheadrightarrow t)$ also has the finite jumps property. Hence, the result now follows by applying Lemma 6.11 to $\mathcal{U}/(s \twoheadrightarrow t)$. □

Lemma 6.15. *If $\mathcal{U}$ has a complete development and if u is a redex in s, then $\mathcal{U} \cup \{u\}$ has a complete development.*

Proof. Perform a complete development of $\mathcal{U}$, resulting in a term t. By definition of valuations and substitutions, nestings of u can only be created by the redexes above u in the initial term. Since there are only finitely many such redexes and since the right-hand side of each rewrite rules satisfies the finite chains condition, only finite chains of residuals of u occur in t (though infinite nestings are still possible). Repeatedly contract a residual of u that is at minimal depth. Since only finite chains of residuals of u occur and since residuals of u cannot nest other residuals of u, it follows by terms being finitely branching that the minimal depth at which redexes are contracted increases after a finite number of steps. Hence, the reduction defined in this way is strongly convergent and contracts all residuals of u, resulting in a complete development of $\mathcal{U} \cup \{u\}$. □

Proposition 6.16. *Let $\mathcal{U}$ have a complete development $s \Rightarrow t$ and let v be a redex in s. The following diagram exists:*

$$\begin{array}{ccc} s & \xrightarrow{v} & t' \\ \Big\Downarrow{\scriptstyle \mathcal{U}} & & \Big\Downarrow{\scriptstyle \mathcal{U}/(s \to t')} \\ t & \underset{v/(s \twoheadrightarrow t)}{\Longrightarrow} & s' \end{array}$$

Proof. Immediate by Lemmas 6.14 and 6.15, Theorem 6.12 and the fact that $(\mathcal{U} \cup \{v\})/(s \to t') = \mathcal{U}/(s \to t')$ and $(\mathcal{U} \cup \{v\})/(s \twoheadrightarrow t) = v/(s \twoheadrightarrow t)$. □

Lemma 6.17. *If $\mathcal{U}$ is finite, then it has a finite complete development.*

Proof. By induction on the number of redexes in $\mathcal{U}$. If $\mathcal{U}$ is empty, we are done. Otherwise, by the finiteness of $\mathcal{U}$, there exists a redex $v \in \mathcal{U}$ such that no redexes from $\mathcal{U}$ occur in its arguments. Contract v. Since no redexes occur in the arguments of v, the set $\mathcal{U}/v$ contains one redex less than $\mathcal{U}$. The induction hypothesis now furnishes the result. □

The following proposition establishes commutativity of a diagram which we will later use to establish certain 'emaciated' projections in Section 8.

Proposition 6.18. *Let $\mathcal{U}$ and $\mathcal{V}$ be sets of redexes in s such that $\mathcal{U}$ has a complete development $s \Rightarrow t$ and $\mathcal{V}$ is finite. The following diagram exists:*

$$\begin{array}{ccc} s & \overset{\mathcal{V}}{\Longrightarrow} & t' \\ \Big\Downarrow{\scriptstyle \mathcal{U}} & & \Big\Downarrow{\scriptstyle \mathcal{U}/(s \twoheadrightarrow t')} \\ t & \underset{\mathcal{V}/(s \twoheadrightarrow t)}{\Longrightarrow} & s' \end{array}$$

Proof. By Lemma 6.17, we have that $\mathcal{V}$ has a finite complete development. Denote this development by

$$s = s_0 \to s_1 \to \cdots \to s_n = t' ,$$

where $s_i \to s_{i+1}$ is assumed to contract a redex v_{i+1}. By Proposition 6.16 we can erect the following diagram, where S_i denotes $s_0 \to^* s_i$ and T_{i+1} denotes $s_i \twoheadrightarrow t_i$ and where $(\mathcal{U}/S_i)/(s_i \to s_{i+1}) = \mathcal{U}/S_{i+1}$ by definition of residuals:

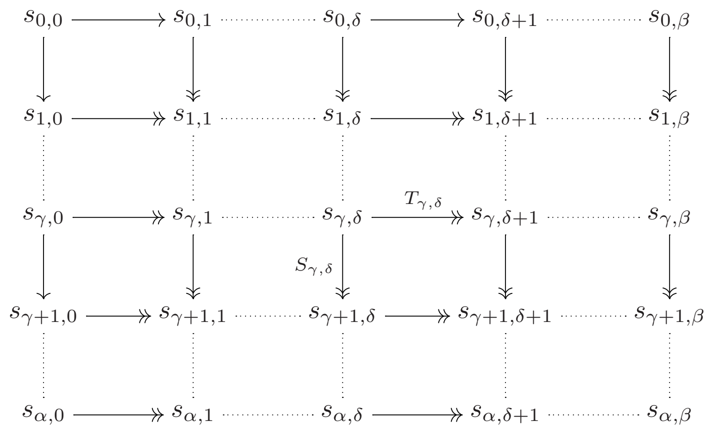
$$\begin{array}{ccccccc} s_0 & \xrightarrow{v_1} & s_1 & \xrightarrow{v_2} & \cdot \cdots\cdots \cdot & \xrightarrow{v_n} & s_n \\ \Big\Downarrow{\scriptstyle \mathcal{U}} & & \Big\Downarrow{\scriptstyle \mathcal{U}/S_1} & & \Big\Downarrow{\scriptstyle \mathcal{U}/S_2} & & \Big\Downarrow{\scriptstyle \mathcal{U}/S_n} \\ t_0 & \underset{v_1/T_1}{\Longrightarrow} & t_1 & \underset{v_2/T_2}{\Longrightarrow} & \cdot \cdots\cdots \cdot & \underset{v_n/T_n}{\Longrightarrow} & t_n \end{array}$$

The reduction $v_1/T_1; v_2/T_2; \ldots ; v_n/T_n$ is a complete development of $\mathcal{V}/T_1$: By definition only residuals from redexes in $\mathcal{V}$ are contracted. Moreover, if not all residuals were contracted, then neither does S_n contract all residuals of $\mathcal{V}$, which is impossible by definition. Hence, by defining $t = t_0$ and $s' = t_n$, the result follows. □

7 Tiling Diagrams

We use the notion of a *tiling diagram* from [17]:

Definition 7.1. A *tiling diagram* of two strongly convergent reductions $S : s_{0,0} \to^{\alpha} s_{\alpha,0}$ and $T : s_{0,0} \to^{\beta} s_{0,\beta}$ is a rectangular arrangement of strongly convergent reductions:



such that (1) each reduction $S_{\gamma,\delta} : s_{\gamma,\delta} \twoheadrightarrow s_{\gamma+1,\delta}$ is a complete development of a set of redexes of $s_{\gamma,\delta}$, and similarly for $T_{\gamma,\delta} : s_{\gamma,\delta} \twoheadrightarrow s_{\gamma,\delta+1}$, (2) the leftmost vertical reduction is S and the topmost horizontal reduction is T, and (3) for each γ and δ the set of redexes developed in $S_{\gamma,\delta}$ is the set of residuals of the redex contracted in $s_{\gamma,0} \to s_{\gamma+1,0}$ across the (strongly convergent) reduction $T_{\gamma,[0,\delta]} : s_{\gamma,0} \to s_{\gamma,1} \to \cdots s_{\gamma,\delta}$ (symmetrically for $T_{\gamma,\delta}$).

For $S_{[0,\alpha],\beta}$ we usually write S/T and we call this reduction the *projection* of S across T (similarly for $T_{\alpha,[0,\beta]}$ and T/S). Moreover, if T consists of a single step contracting a redex u, we also write S/u (symmetrically T/u).

The following extends Theorem 12.6.5 in [17], where it is assumed that S and T are reductions of limit ordinal length, to reductions of arbitrary length:

Theorem 7.2. *Let S and T be strongly convergent reductions starting from the same term. Suppose that the tiling diagram for S and T exists except that it is unknown if S/T and T/S are strongly convergent and end in the same term. The following are equivalent:*

1. *The tiling diagram of S and T can be completed, i.e. S/T and T/S are strongly convergent and end in the same term.*

2. *S/T is strongly convergent.*

3. *T/S is strongly convergent.*

Proof. Obviously, the first statement trivially implies the second and third. Hence, we only need to prove that the first holds under assumption of either the second or the third statement. Without loss of generality there are three cases to consider depending on the lengths of S and T, i.e. α and β.

- In case $\alpha = \alpha' + 1$ and $\beta = \beta' + 1$, write $\mathcal{U} = u/T_{\alpha',[0,\beta']}$ and $\mathcal{V} = v/S_{[0,\alpha'],\beta'}$, where u is the redex contracted in $s_{\alpha',0} \to s_{\alpha,0}$ and v the redex contracted in $s_{0,\beta'} \to s_{0,\beta}$.

  Assume that S/T is strongly convergent. By definition of tiling diagrams, $S_{\alpha',\beta'}$ is a (not necessarily complete) development of $\mathcal{U} \cup \mathcal{V}$ and $T_{\alpha',\beta'}; S_{\alpha',\beta}$ is a complete development of $\mathcal{U} \cup \mathcal{V}$. By Lemma 6.14 and Theorem 6.12, it now follows that there exists a complete development of $(\mathcal{U} \cup \mathcal{V})/S_{\alpha',\beta'}$ that starts in $s_{\alpha',\beta}$ and ends in the same term as $T_{\alpha',\beta'}; S_{\alpha',\beta}$. Since $(\mathcal{U} \cup \mathcal{V})/S_{\alpha',\beta'} = \mathcal{V}/S_{\alpha',\beta'}$ by definition of $S_{\alpha',\beta'}$, the complete development is in fact a complete development of residuals of v in $s_{\alpha',\beta}$. Hence, it completes the tiling diagram as required.

  Assume now that T/S is strongly convergent. Since α and β are both successor ordinals the proof proceeds completely symmetrical to the case in which strong convergence of S/T is assumed.

- In case α is a limit ordinal and $\beta = \beta' + 1$, write v for the redex contracted in $s_{0,\beta'} \to s_{0,\beta}$. That S/T and T/S end in the same term follows immediately in case S/T and T/S are both strongly convergent, since this implies that for larger $\gamma < \alpha$ more residuals of v up to greater depths must occur at the same positions in $s_{\gamma,\beta'}$ and $s_{\alpha,\beta'}$. Hence, we only need to prove strong convergence.

  Assume that S/T is strongly convergent and T/S is not. By assumption there exists a position p of minimal depth d such that an infinite number of residuals of v are contracted in T/S. That is, since residuals of v cannot nest each other, an infinite collapsing chain of residuals of v exists in $s_{\alpha,\beta'}$. By strong convergence of $S_{[0,\alpha],\beta'}$, there exists a $\gamma < \alpha$ such that all steps in $S_{[\gamma,\alpha],\beta'}$ occur below depth d. By definition of γ and the infinite chain in $s_{\alpha,\beta'}$, each $s_{\kappa,\beta}$ with $\gamma < \kappa < \alpha$ has a finite collapsing chain of residuals of v at position p. No infinite chain can occur at position p in $s_{\kappa,\beta'}$, since each $T_{\kappa,\beta'}$ is strongly convergent. The finite chains of residuals of v become arbitrary large along $S_{[\gamma,\alpha],\beta'}$, otherwise no infinite chain exists in $s_{\alpha,\beta'}$. However, this implies that for every point along the strongly convergent reduction S/T a redex is contracted at position p somewhere later along the reduction, contradiction. Hence, T/S is strongly convergent.

  Assume now that T/S is strongly convergent and S/T is not. By assumption there exists a position p of minimal depth d such that an infinite number of reductions occur at p in S/T. Moreover, by strong convergence of $S_{[0,\alpha],\beta'}$ and the minimality of d, there exists a $\gamma < \alpha$ such that all redexes contracted in $S_{[\gamma,\alpha],\beta'}$ and $S_{[\gamma,\alpha],\beta}$ occur at depth d or below.

  In $s_{\gamma,\beta'}$ only a finite number of residuals of v occur at depth less than d. Hence, the subterms of $s_{\gamma,\beta}$ at depth d or below consist of finite chains of parallel subterms that occur at depth d in $s_{\gamma,\beta'}$, with all residuals of

v contracted. By definition of γ and since the chains consist of parallel subterms, it follows that all redexes contracted in $S_{[\gamma,\alpha],\beta}$ occur within the chains and that no further nestings can be created among the chains. In fact, since the subterms are parallel, further nestings cannot even be created within the chains. But then, there exists a point along $S_{[\gamma,\alpha],\beta}$ such that precisely one of the subterms of $s_{\gamma,\beta'}$ is responsible for the infinite number of reductions at p. Since S is strongly convergent, this implies that the subterm has at its root a collapsing chain of residuals of v that becomes arbitrary large along $S_{[\gamma,\alpha],\beta'}$. Whence, there exists an infinite collapsing chain in the limit $s_{\alpha,\beta'}$. Since the chain cannot be erased by contracting other residuals of v, otherwise the infinite reduction in at p in $S_{[\gamma,\alpha],\beta}$ does not exist, it follows that T/S cannot be strongly convergent as a complete development of residuals of v in $s_{\alpha,\beta'}$, contradiction. Hence, S/T is strongly convergent.

- If α and β are limit ordinals, then the result follows by Theorem 12.6.5 in [17]. The proof is independent of the details of rewriting, except for its use of one lemma — Lemma 12.5.12 — which also holds in the case of iCRSs and is Lemma 4.18 above. □

8 Essentiality

Considering only *fully-extended*, *orthogonal* iCRSs from this section onwards, we define a 'measure' on finite sequences of complete developments. The measure satisfies certain favourable properties with respect to projection of reductions and is employed in Sections 9 and 10. The measure is inspired by a proof technique originally developed by Sekar and Ramakrishnan [32] to study normalising strategies in rewriting. The technique was later refined by Middeldorp [27] and extended to higher-order rewriting by Van Oostrom [38]. We build on the latter work. Where the techniques of Sekar and Ramakrishnan, Middeldorp, and Van Oostrom apply to finite reductions, ours applies to finite sequences of complete developments. A shift to finite sequences of complete developments is necessary, because in the current setting projecting one reduction step over another may yield an infinite complete development of the residuals of the projected redex.

8.1 Prefixes

To define the measure on finite sequences of complete developments, we first define the notion of *(path) prefix* and those of *essential* positions and redexes.

Definition 8.1. A *prefix* of a term s is a finite set $P \subseteq \mathcal{P}os(s)$ such that all prefixes of positions in P are also in P.

Take heed that prefixes are *finite*!

We now relate paths, as defined in Section 6.1, with prefixes. In particular, we need a notion of paths that 'occur' in a prefix of a term where some set of redexes is present.

Definition 8.2. Let s and t be terms, $\mathcal{U}$ a set of redexes in s such that $s \Rightarrow^{\mathcal{U}} t$, and P a prefix of t. The *path prefix* of P with respect to $\mathcal{U}$ is the set of all paths Π of s with respect to $\mathcal{U}$ such that the concatenation of the edge labels of the path projection $\phi(\Pi)$ is in P.

Example 8.3. Consider the iCRS of Example 6.5 and the terms $s = f([x]g(x), a)$ and $t = g(g(g(a)))$. Observe that $s \to t$. The set $P = \{\epsilon, 1, 11\}$ is a prefix of t. Let $\mathcal{U}$ be the set containing the only redex of s. The path prefix of P with respect to $\mathcal{U}$ is the set of all paths that are prefixes of

$$(s, \epsilon) \to (r, \epsilon, \epsilon) \to (s, 10) \xrightarrow{1} (s, 101) \to (r, 1, \epsilon) \xrightarrow{1} (r, 11, \epsilon) \to (s, 10)\,.$$

Given a path prefix with respect to some set of redexes in a term, we would like to recover the positions 'encountered' by the paths in the path prefix, in particular the positions in the redex patterns encountered by the paths. The following map facilitates this recovery:

Definition 8.4. Let s be a term and $\mathcal{U}$ a set of redexes in s. The map ζ from *finite* paths Π of s with respect to $\mathcal{U}$, with final node n, to *finite* subsets of $\mathcal{P}os(s)$ is defined as follows:

$$\zeta(\Pi) = \begin{cases} \{p\} & \text{if } n = (s, p) \text{ and no redex in } \mathcal{U} \text{ occurs at } p \\ Q & \text{if } n = (s, p) \text{ and a redex } u \in \mathcal{U} \text{ occurs at } p \\ \emptyset & \text{if } n = (r, p, p_u) \end{cases}$$

where Q is the set of positions of s that occur in the redex pattern of u.

The following lemma shows that ζ can be extended to a well-defined function on path prefixes:

Lemma 8.5. *Let s and t be terms, $\mathcal{U}$ a set of redexes in s such that $s \Rightarrow^{\mathcal{U}} t$, and P a prefix of t. If Ψ is the path prefix of P with respect to $\mathcal{U}$, then $\zeta(\Psi) = \{\zeta(\Pi) \mid \Pi \in \Psi\}$ is well-defined and a prefix of s.*

Proof. Let Ψ be the path prefix of P with respect to $\mathcal{U}$. Since $\mathcal{U}$ has a complete development, it follows by Lemma 6.11 that $\mathcal{U}$ also has the finite jumps property, i.e. all path projections in $\mathcal{P}(s,\mathcal{U})$ contain only finite sequences of unlabelled nodes and ϵ-labelled edges. As each path is a prefix of a maximal path, whose path projections are in $\mathcal{P}(s,\mathcal{U})$, it follows by definition of path projections and the finite jumps property that each path in Ψ is finite. Hence, $\zeta(\Psi)$ is well-defined.

For each position in the prefix P of t a finite number of paths is included in Ψ. This follows by induction on the length of the positions, employing the fact that all path projections contain only finite sequences of unlabelled nodes and ϵ-labelled edges and the fact that the extension of a path is uniquely determined by definition of paths and the considered position. Since P is finite, the same follows for Ψ. Hence, as ζ maps each finite path to a finite number of positions, $\zeta(\Psi)$ is finite.

Suppose for a certain $p \in \zeta(\Psi)$ that there exists a $q < p$ such that $q \notin \zeta(\Psi)$. There are two possibilities: q occurs either in the redex pattern of a redex in $\mathcal{U}$, or not. In case q occurs in the redex pattern of a redex $u \in \mathcal{U}$, it follows by $p \in \zeta(\Psi)$ and the definition of paths that there exists a path in the path prefix which ends in the node (s, p_u), with p_u the position of the redex u. However, we then have that $q \in \zeta(\Psi)$ by definition of ζ, contradiction. In case q does not occur in a redex pattern, it follows by the definition of paths and the inclusion of p in $\zeta(\Psi)$ that there is a path in the path prefix which ends in the node (s, q), again a contradiction by definition of ζ. Hence, all prefixes of positions in $\zeta(\Psi)$ are included in $\zeta(\Psi)$. Employing the finiteness of $\zeta(\Psi)$ the result now follows. □

By the previous lemma, it is easy to see that the following is well-defined:

Definition 8.6. Let s and t be terms, $\mathcal{U}$ a set of redexes in s such that $s \Rightarrow^{\mathcal{U}} t$, and P a prefix of t. A position $p \in \mathcal{P}os(s)$ is called *essential* for P if $p \in \zeta(\Psi)$ with Ψ the path prefix of P with respect to $\mathcal{U}$.

Thus, intuitively, a position is essential if it 'contributes' to the prefix of the final term t of a complete development.

Example 8.7. In Example 8.3, consider the prefix P. The positions ϵ, 1, 10, and 101 are essential for P in s.

The next proposition shows that an essential position will always descend to a position in the assumed prefix in case the position does not occur in the redex pattern of any redex in the assumed complete development.

Proposition 8.8. *Let s and t be terms, $\mathcal{U}$ a set of redexes in s such that $s \Rightarrow^{\mathcal{U}} t$, and P a prefix of t. If $p \in \mathcal{P}os(s)$ does not occur in the redex pattern of any redex in $\mathcal{U}$ and is not the position of a variable bound by a redex in $\mathcal{U}$, then p is essential iff there exists a $q \in P$ such that $q \in p/(s \Rightarrow t)$ and p is inessential iff no descendant of p occurs in P.*

Proof. By Lemma 6.11, it follows that $\mathcal{U}$ has the finite jumps property. Employing the labelling from the proof of Theorem 6.10(2) and its properties, it is easy to see that a position $p \in \mathcal{P}os(s)$ descents to a position $q \in \mathcal{P}os(t)$ iff p does not occur in a redex pattern of a redex in $\mathcal{U}$ and there exists a finite path Π with final node $n = (s, p)$ such that $\phi(n)$ is labelled and such that the concatenation of the edge labels of the path projection of Π is q. Since $\zeta(\Pi) = \{p\}$, the result follows by definition of path prefixes. □

As the set of positions obtained through application of ζ is a prefix by Lemma 8.5, essentiality is easily extended to a finite sequence of complete developments $s_0 \Rightarrow^{\mathcal{U}_1} s_1 \Rightarrow^{\mathcal{U}_2} \cdots \Rightarrow^{\mathcal{U}_n} s_n$: In case of s_n, define the positions of P as essential. In case of s_i, with $i < n$, define the positions that are essential for the essential positions of s_{i+1} as essential. We can now define the following:

Definition 8.9. A redex in a term s_i along $s_0 \Rightarrow^{\mathcal{U}_1} s_1 \Rightarrow^{\mathcal{U}_2} \cdots \Rightarrow^{\mathcal{U}_n} s_n$ is called *essential* for a prefix P of s_n, if a residual of the redex occurs in a set $\mathcal{U}_j$ with

$j > i$ and if the position at which the residual occurs is essential. The redex is called *inessential* otherwise.

8.2 Measure

In this section, we assume that D denotes a finite sequence of complete developments $s_0 \Rightarrow^{\mathcal{U}_1} s_1 \Rightarrow^{\mathcal{U}_2} \cdots \Rightarrow^{\mathcal{U}_n} s_n$. Moreover, we assume that P is a prefix of s_n, the final term of D. We define a measure on D with respect to P:

Definition 8.10. The *measure* $\mu_P(D)$ of the finite sequence of complete developments D with respect to P is the n-tuple $(l_n, \ldots, l_1)$ — note the reverse order! — such that l_i, with $1 \leq i \leq n$, denotes the cardinality of the path prefix of P_i with respect to $\mathcal{U}_i$, where P_i is the set of positions in s_i essential for P.

The tuples in the above definition are compared first length-based and then lexicographically (in the natural order). This yields a well-founded order, as each element of a tuple is finite by Lemma 8.5. We denote the order by $\prec$.

Before proving any properties of the measure, we introduce some useful terminology:

Definition 8.11. Let s and t be terms and P a prefix of s. The term t *mirrors* s in P, if $P \subseteq \mathcal{P}os(t)$ and $root(t|_p) = root(s|_p)$ for all $p \in P$.

Definition 8.12. Let D' denote $t_0 \Rightarrow^{\mathcal{V}_1} t_1 \Rightarrow^{\mathcal{V}_2} \cdots \Rightarrow^{\mathcal{V}_n} t_n$. If t_n mirrors s_n in the prefix P, then D' *mirrors* D in P if for all $0 \leq i \leq n$ it holds that t_i mirrors s_i in P_i and P_i is the set positions essential for P in both s_i and t_i.

The following lemma is key for the use of the measure:

Lemma 8.13. *For D there exists a finite sequence of complete developments $D' : t_0 \Rightarrow^{\mathcal{V}_1} t_1 \Rightarrow^{\mathcal{V}_2} \cdots \Rightarrow^{\mathcal{V}_n} t_n$, with $\mathcal{V}_i$ consisting of a* finite *number of* essential *redexes for all $1 \leq i \leq n$, such that D' mirrors D in P and $\mu_P(D') = \mu_P(D)$.*

Proof. By induction on n, the number of complete developments in D. In case $n = 0$, define $t_0 = s_0$ and the result is immediate by definition of t_0.

In case $n > 0$, let $\mathcal{U}'_n$ contain the redexes from $\mathcal{U}_n$ essential for P and write P' for the set of positions of s_{n-1} essential for P. Observe for each $u \in \mathcal{U}'_n$ that all positions in the redex pattern of u occur at positions in P' by definition of ζ. Hence, since we have by the induction hypothesis that t_{n-1} mirrors s_{n-1} in P', it follows by orthogonality and fully-extendedness that there exists for each redex in $\mathcal{U}'_n$ a redex in t_{n-1} such the positions and the employed redex patterns are identical. Define $\mathcal{V}_n$ to be the set of these corresponding redexes in t_{n-1}. Obviously, the sets $\mathcal{V}_n$ and $\mathcal{U}'_n$ have the same cardinality, which is finite because P' is finite.

Since $\mathcal{V}_n$ is finite, it follows by Lemma 6.17 that there exists a complete development $t_{n-1} \Rightarrow^{\mathcal{V}_n} t_n$. Moreover, since P' is a prefix and t_{n-1} mirrors s_{n-1} in P', it follows by definition of paths and $\mathcal{V}_n$ that for each path of s_{n-1} with respect to $\mathcal{U}_n$ occurring in the path prefix of P there exists an identical path of t_{n-1} with respect to $\mathcal{V}_n$. Hence, by definition of path projections, we have for the terms matching $\mathcal{P}(s_{n-1},\mathcal{U}_n)$ and $\mathcal{P}(t_{n-1},\mathcal{V}_n)$, i.e. s_n and t_n, that $P \subseteq \mathcal{P}os(t_n)$, $root(s_n|_p) = root(t_n|_p)$ for all $p \in P$, and that all positions in P' and redexes in $\mathcal{V}_n$ are essential for P. The induction hypothesis now furnishes the result. □

Observe that the above lemma 'cuts down' the sets of redexes that occur in the sequence of complete developments to *finite* sets consisting solely of essential redexes. The lemma states that this suffices to obtain a term t_n with prefix P.

We can now define the following:

Definition 8.14. Let $\mathcal{U}_i$ be finite for all $1 \leq i \leq n$ in D. If $s_0 \to t_0$ contracts a redex u such that no redex in u/D occurs at a position in P, then the *emaciated projection* of D across $s_0 \to t_0$ with respect to P, written $D/\!\!/u$, is defined by application of Lemma 8.13 to D/u.

That the projection D/u in the above definition exists follows by repeated application of Proposition 6.18. The final term of D/u mirrors the final one of D in P, as no redex of u/D occurs at a position in P. Hence, Lemma 8.13 can be applied and the final term of $D/\!\!/u$ mirrors the final one of D in P.

The next two lemmas concern changes to the measure when taking emaciated projections.

Lemma 8.15. *Let $\mathcal{U}_i$ be finite for all $1 \leq i \leq n$ in D. If $s_0 \to t_0$ contracts an inessential redex u such that no redex in u/D occurs at a position in P, then the position of u is inessential for P, $D/\!\!/u$ mirrors D in P, and $\mu_P(D/\!\!/u) = \mu_P(D)$.*

Proof. Suppose that $s_0 \to t_0$ contracts an inessential redex u such that no redex in u/D occurs at a position in P. Denote D/u by $t_0 \Rightarrow^{\mathcal{V}_1} t_1 \Rightarrow^{\mathcal{V}_2} \cdots \Rightarrow^{\mathcal{V}_n} t_n$ where $\mathcal{V}_i = \mathcal{U}_i/(s_{i-1} \Rightarrow t_{i-1})$ for all $1 \leq i \leq n$.

For all $0 \leq i < n$, no residual of u occurs at an essential position in s_i. Otherwise, a residual also occurs at an essential position in s_{i+1} by definition of residuals and Proposition 8.8. Iteratively, a residual then occurs in s_n at a position in P, contradicting assumptions. Hence, by induction we have that t_i mirrors s_i in P_i, where P_i is the set of position essential for P in both s_i and t_i. Moreover, for each essential redex in $\mathcal{U}_j$, with $1 \leq j \leq n$, there exists an essential redex in $\mathcal{V}_j$ such that the redex positions and the employed redex patterns are identical, and vice versa.

Write $\mu_P(D) = (l_n, \ldots, l_1)$ and $\mu_P(D/u) = (l'_n, \ldots, l'_1)$. For all $1 \leq j \leq n$ the path prefixes of $\mathcal{V}_j$ and $\mathcal{U}_j$ have the same cardinality by the correspondence between the essential redexes, i.e. $l'_j = l_j$. Thus, $\mu_P(D/u) = \mu_P(D)$ and by Lemma 8.13 it follows that $\mu_P(D/\!\!/u) = \mu_P(D)$. □

Lemma 8.16. *Let $\mathcal{U}_i$ be finite for all $1 \leq i \leq n$ in D. If $s_0 \to t_0$ contracts an essential redex u such that no redex in u/D occurs at a position in P, then $\mu_P(D/\!\!/u) \prec \mu_P(D)$.*

Proof. Suppose that $s_0 \to t_0$ contracts an essential redex u such that no redex in u/D occurs at a position in P. Denote D/u by $t_0 \Rightarrow^{\mathcal{V}_1} t_1 \Rightarrow^{\mathcal{V}_2} \cdots \Rightarrow^{\mathcal{V}_n} t_n$ where $\mathcal{V}_i = \mathcal{U}_i/(s_{i-1} \Rightarrow t_{i-1})$ for all $1 \leq i \leq n$.

Let i be the largest index of a set $\mathcal{U}_i$ such that the set contains a residual of u that is essential. No residual of u occurs at an essential position in s_j with $i < j < n$. Otherwise, a residual also occurs at an essential position in s_{j+1} by definition of residuals and Proposition 8.8. Iteratively, a residual then occurs in s_n at a position in P, contradicting assumptions. Hence, by induction we have for all $i < j \leq n$ that t_j mirrors s_j in P_j, where P_j is the set of positions essential for P in both s_j and t_j. Moreover, for each essential redex in $\mathcal{U}_j$, there exists an essential redex in $\mathcal{V}_j$ such that the redex positions and the employed redex patterns are identical, and vice versa.

Write $\mu_P(D) = (l_n, \ldots, l_1)$ and $\mu_P(D/u) = (l'_n, \ldots, l'_1)$. For all $j < i$, the cardinality of the path prefix of $\mathcal{V}_j$ may be greater than the one of $\mathcal{U}_j$, i.e. we may have $l'_j \neq l_j$. The cardinality of the path prefix of $\mathcal{V}_i$ is less than that of $\mathcal{U}_i$ by Proposition 6.6 and Lemma 6.7. Hence, $l'_i < l_i$. Finally, for all $i < k \leq n$ the path prefixes of $\mathcal{V}_k$ and $\mathcal{U}_k$ have the same cardinality by the correspondence between the essential redexes, i.e. $l'_k = l_k$. Thus, $\mu_P(D/u) \prec \mu_P(D)$ and by Lemma 8.13 it follows that $\mu_P(D /\!\!/ u) \prec \mu_P(D)$. □

The final lemma of this section allows us to 'simulate' a development of redexes in a term s by another development of redexes in term mirroring s without changing the measure of the development.

Lemma 8.17. *Let $\mathcal{U}_i$ be finite for all $1 \leq i \leq n$ in D. If a term t_0 mirrors s_0 in the positions of s_0 essential for P, then there exists a finite sequence of complete developments D' starting in t_0 such that D' mirrors D in P and $\mu_P(D') = \mu_P(D)$.*

Proof. By Lemma 8.13 we may assume for each $1 \leq i \leq n$ that all redexes in $\mathcal{U}_i$ are essential. The proof proceeds by induction on n, the number of complete developments in D. In case $n = 0$, the result is immediate by assumption.

In case $n > 0$, it follows by the induction hypothesis that t_{n-1} mirrors s_{n-1} in P', with P' the set of positions of s_{n-1} essential for P. Hence, by orthogonality and fully-extendedness there exists for each redex in $\mathcal{U}_n$ a redex in t_{n-1} such that the positions and the employed redex patterns are identical. Define $\mathcal{V}_n$ to be the set of these corresponding redexes in t_{n-1}. Obviously, the sets $\mathcal{V}_n$ and $\mathcal{U}_n$ have the same cardinality, which is finite because P' is finite.

Since P' is a prefix and t_{n-1} mirrors s_{n-1} in P', it follows by definition of paths and $\mathcal{V}_n$ that for each path of s_{n-1} with respect to $\mathcal{U}_n$ occurring in the path prefix of P there exists an identical path of t_{n-1} with respect to $\mathcal{V}_n$. Hence, by definition of path projections, we have for the terms matching $\mathcal{P}(s_{n-1}, \mathcal{U}_n)$ and $\mathcal{P}(t_{n-1}, \mathcal{V}_n)$, i.e. s_n and t_n, that $P \subseteq \mathcal{P}os(t_n)$, $root(s_n|_p) = root(t_n|_p)$ for all $p \in P$, and that all positions in P' and redexes in $\mathcal{V}_n$ are essential for P. The induction hypothesis now furnishes the result. □

9 Confluence

It is well-known that confluence does not hold for iTRSs even under assumption of orthogonality [18]. Since each iTRS can be viewed as a fully-extended iCRS, it follows that fully-extended, orthogonal iCRSs are in general not confluent either. In the case of iTRSs two approaches are known for restoring confluence [18]: identifying all subterms that disrupt confluence and restricting the rules that are allowed. Identifying all subterms that disrupt confluence leads to the definition of so-called hypercollapsing subterms and the result that orthogonal iTRSs are confluent modulo these subterms. Restricting the rules that are allowed leads to results regarding almost non-collapsing iTRSs.

Continuing to consider only *fully-extended*, *orthogonal* iCRSs, we next prove that these iCRSs are confluent modulo hypercollapsing subterms, where a term s is called hypercollapsing if for every $s \twoheadrightarrow t$ we have that t is reducible to a collapsing redex. The result generalises similar results for iTRSs and iλc. Alas, the proofs for iTRSs and iλc from [17] cannot be lifted to the general higher-order case: For iTRSs the proof hinges on the Strip Lemma and for iλc it hinges on the notion of head reduction, both of which fail to properly generalise to iCRSs. To circumvent these problems, we employ the measure defined in the previous section.

Apart from confluence modulo, we show in Section 9.3 that the positive result that an iTRS is confluent iff it is almost non-collapsing cannot be trivially lifted to iCRSs.

Remark 9.1. On a historical note: Courcelle [6] notes similar problems with confluence while trying to define second-order substitutions on infinite trees. He works around these problems by requiring rules to be non-collapsing. In a general setting such as ours this would be too harsh of a restriction.

9.1 Hypercollapsingness

We now treat a special kind of troublesome reductions and terms.

Definition 9.2. A *hypercollapsing reduction* of length α is an open strongly continuous reduction with an infinite number of root-collapsing steps.

Thus, a hypercollapsing reduction is a transfinite reduction of limit ordinal length which is not strongly *convergent*, as such the term s_α is omitted. Note that if we write $(s_\beta)_{\beta<\alpha}$ for a hypercollapsing reduction sequence, then we have that every initial sequence $(s_\beta)_{\beta<\gamma+1}$ with $\gamma < \alpha$ is strongly convergent.

Example 9.3. Hypercollapsing reductions are known even in the first order case where we have, e.g. (in the syntax of iCRSs) the reduction rule $f(Z) \to Z$ and the term f^ω from which there is the hypercollapsing reduction

$$f^\omega \to f^\omega \to \cdots$$

which is obtained by repeatedly contracting the redex at the root.

For example of a more higher-order spirit consider the rule $g([x]Z(x)) \to Z([x]Z(x))$. From the term $g([x]g(x))$ there is the hypercollapsing reduction

$$g([x]g(x)) \to g([x]g(x)) \to \cdots .$$

The crucial definition is now the following:

Definition 9.4. A term s is said to be *hypercollapsing* if, for all terms t with $s \twoheadrightarrow t$, there exists a term t' with $t \twoheadrightarrow t'$ such that t' has a collapsing redex at the root.

It is not hard to see that a hypercollapsing term has a hypercollapsing reduction starting from it; the converse, however, is much more difficult, and is contained in the following lemma, to the proof of which we devote the remainder of the section.

Lemma 9.5. *Let s be a term. If there is a hypercollapsing reduction starting from s, then s is hypercollapsing.*

To start, we observe that hypercollapsing reductions satisfy a 'compression' property:

Lemma 9.6. *Let s be a term. If there is a hypercollapsing reduction starting from s, then there is a hypercollapsing reduction of length ω starting from it.*

Proof. By definition of hypercollapsing reductions we may write a hypercollapsing reduction starting from s as:

$$s = s_0 \twoheadrightarrow s'_0 \to s_1 \twoheadrightarrow s'_1 \to s_2 \twoheadrightarrow \cdots ,$$

where for all $i \in \mathbb{N}$ we have that $s'_i \to s_{i+1}$ is root-collapsing and such that no root-collapsing steps occur in $s_i \twoheadrightarrow s'_i$.

To show that a hypercollapsing reduction of length ω exists, let $i \in \mathbb{N}$ be arbitrary and assume there exists a term t_i such that $t_i \twoheadrightarrow^{\leq\omega} s_i$. By definition, the reduction $t_i \twoheadrightarrow s_i \twoheadrightarrow s'_i \to s_{i+1} \twoheadrightarrow \cdots$ is hypercollapsing and by compression we have $t_i \to^* t'_i \to t_{i+1} \twoheadrightarrow^{\omega} s_{i+1}$ with $t'_i \to t_{i+1}$ a root-collapsing step. Hence, if we define $t_0 = s_0 = s$, we obtain the following reduction:

$$s = t_0 \to^* t'_0 \to t_1 \to^* t'_1 \to t_2 \to^* \cdots ,$$

where for all $i \in \mathbb{N}$ we have that $t'_i \to t_{i+1}$ is root-collapsing and such that no root-collapsing steps occur in $t_i \to^* t'_i$. Since each $t_i \to^* t'_i$ is finite and is followed by a root-collapsing step $t'_i \to t_{i+1}$, there exists a hypercollapsing reduction of length ω starting from s, as required. $\square$

The following is the iCRS analogue of Lemma 12.8.4 in [17] for iTRSs and strengthening for iλc:

Lemma 9.7. *Let s be a term. If there exists a hypercollapsing reduction starting from s, and a rewrite step $s \to t$, then there is a hypercollapsing reduction starting from t.*

Proof. Define $s_0 = s$, $t_0 = t$, and suppose that u is the redex contracted in $s \to t$. By Lemma 9.6, we may write the hypercollapsing reduction starting in s_0 as:

$$s_0 \to^* s'_0 \to s_1 \to^* s'_1 \to s_2 \to^* \cdots ,$$

where for all $i \in \mathbb{N}$, $s'_i \to s_{i+1}$ is root-collapsing and such that no root-collapsing steps occur in $s_i \to^* s'_i$. By repeated application of Proposition 6.16 we can erect the following diagram:

$$\begin{array}{ccccccccccc} s_0 & \xrightarrow{*} & s'_0 & \longrightarrow & s_1 & \xrightarrow{*} & s'_1 & \longrightarrow & s_2 & \xrightarrow{*} & \cdot \cdots \\ \downarrow{\scriptstyle u} & & \Downarrow{\scriptstyle \mathcal{U}'_0} & & \Downarrow{\scriptstyle \mathcal{U}_1} & & \Downarrow{\scriptstyle \mathcal{U}'_1} & & \Downarrow{\scriptstyle \mathcal{U}_2} & & \vdots \\ t_0 & \twoheadrightarrow & t'_0 & \twoheadrightarrow & t_1 & \twoheadrightarrow & t'_1 & \twoheadrightarrow & t_2 & \twoheadrightarrow & \cdot \cdots \end{array}$$

Write S_i for $s_i \to^* s'_i \to s_{i+1} \twoheadrightarrow \cdots$ and T_i for $t_i \twoheadrightarrow t'_i \twoheadrightarrow t_{i+1} \twoheadrightarrow \cdots$. If it holds for each $i \in \mathbb{N}$ that a root-collapsing step occurs in T_i, then an infinite number of root-collapsing steps occurs in T_0 implying that the reduction is hypercollapsing.

To show that a root-collapsing step occurs in each T_i we distinguish two cases: (1) a root-collapsing step occurs in S_i that does not contract a residual of u, and (2) all root-collapsing steps contract residuals of u. We deal with each of these cases in turn:

1. In this case there is a root-collapsing step $s'_j \to s_{j+1}$ with $j > i$ where the contracted redex, say v, is not a residual of u. Since $\mathcal{U}'_j$ by construction contracts only residuals of u, we have by orthogonality that a residual of v occurs at the root of t'_j and that no other residuals of v occur in t'_j. Also by construction, $t'_j \twoheadrightarrow t_{j+1}$ contracts precisely all residuals of v. Hence, $t'_j \twoheadrightarrow t_{j+1}$ is a root-collapsing step.

2. In this case, first observe that u is a collapsing redex and that no infinite chains of residuals of u can occur in any s_k and s'_k with $k \geq i$ by the fact that only a finite number of reductions occur before s_k and s'_k, and because right-hand sides of rules only allow for finite chains of meta-variables.

   Next observe that all terms in the topmost reduction must have a residual of u at the root, since we would otherwise have a root-step in some term that brings a residual of u to the root. Such a step is by definition root-collapsing.

   Finally, observe that the residual of u at the root of s_i is part of a finite chain of residuals of u, where each residual next in the chain is substituted for the variable that occurs at the root of the right-hand side of the rewrite rule employed in contracting u.

   As residuals of u cannot create further nestings of other residuals of u, we have for each step following s_i that each of the residuals in the chain starting at the root has at most one residual. Eventually the number of residuals must become zero as (1) u is a collapsing redex and as (2) an infinite number of root steps occur in S_i. Since the residuals always occur in a chain starting at the root we have that the last of the residuals is always contracted by means of a root step, say $s'_j \to s_{j+1}$.

   Now suppose that no collapsing redexes are contracted in $s_i \to^* s_{j+1}$ that have a residual occurring at the root of one of the terms in $t_i \twoheadrightarrow t_{j+1}$.

Then, since residuals of u occur in finite chains and cannot nest other residuals of u and since every development of $\mathcal{U}_k$ and $\mathcal{U}'_k$ contracts only residuals of u, it follows that no residual of u can occur at the root of s_{j+1}, a contradiction. Hence, at least one root-collapsing step occurs in $t_i \twoheadrightarrow t_{j+1}$.

As required, we have that a root-collapsing step occurs in each T_i. Hence, T_0 is a hypercollapsing reduction starting from $t_0 = t$. □

The next lemma shows that the property of being reducible to a term with a collapsing redex at the root cannot be destroyed by reductions, unless they contain a collapsing step at the root themselves.

Lemma 9.8. *If $s \twoheadrightarrow t$ contains no root-collapsing steps and s reduces to a collapsing redex, then so does t.*

Proof. We show by ordinal induction that every term s_α in $s \twoheadrightarrow t$ reduces to a collapsing redex by a finite sequence D_α of complete developments, where each set of redexes is finite. Denote by P_α the set of positions of the redex pattern at the root of the final term of D_α. To facilitate the induction we also show for all $\beta \leq \alpha$ that $\mu_{P_\alpha}(D_\alpha) \preceq \mu_{P_\beta}(D_\beta)$ and that s_α mirrors s_β in the positions of s_β essential for P_β in case $\mu_{P_\alpha}(D_\alpha) = \mu_{P_\beta}(D_\beta)$.

For $s_0 = s$, it follows by assumption that s_0 reduces to a root-collapsing redex. In fact, by strong convergence and compression, s_0 reduces to a root-collapsing redex by a finite reduction D_0. As any finite reduction can be seen as a finite sequence of complete developments, where each set of redexes is a singleton set, the result follows.

For $s_{\alpha+1}$, there are two cases to consider with respect to the redex u contracted in $s_\alpha \to s_{\alpha+1}$ depending on the occurrence of a residual of u at the root of the final term of D_α:

- In case no residual of u occurs at the root of the final term of D_α, we discriminate between u being either essential or inessential for P_α. If u is essential, the result follows by Lemma 8.16 and the induction hypothesis. If u is inessential, the result follows by Lemma 8.15 and the induction hypothesis.

- In case a residual of u occurs at the root of the final term of D_α, a root-collapsing step not contracting a residual of u occurs somewhere along D_α. Otherwise, a residual of u cannot occur at the root of the final term of D_α since u itself is not root-collapsing. Hence, there exists a finite sequence D'_α of complete developments, where each set of redexes is finite, that is shorter than D_α and that has a collapsing redex, other than a residual of u, at the root of its final term. By definition of $\prec$, it follows that $\mu_{P'_\alpha}(D'_\alpha) \prec \mu_{P_\alpha}(D_\alpha)$, where P'_α is the set of positions of the redex pattern at the root of the final term of D'_α. The case in which no residual of u occurs at the root of the final term of the complete development now applies and the result follows.

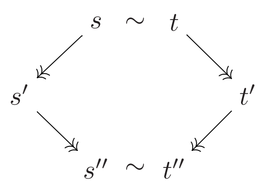


Figure 6: Definition 9.9

For s_α with α a limit ordinal, it follows by the well-foundedness of $\prec$ that there exist a $\beta < \alpha$ such that for every $\beta < \gamma < \alpha$ we have $\mu_{P_\gamma}(D_\gamma) = \mu_{P_\beta}(D_\beta)$. Hence, by the induction hypothesis it follows for all $\beta < \gamma < \alpha$ that s_γ mirrors s_β in the positions of s_β essential for P_β and, by strong convergence, s_α mirrors s_β in these positions. Consider D_β. By Lemma 8.13 there exists a finite sequence of complete developments D'_β starting in s_β contracting only essential redexes such that the final term of D'_β has a collapsing redex at the root and such that $\mu_{P'_\beta}(D'_\beta) = \mu_{P_\beta}(D_\beta)$, where P'_β is the set of positions of the redex pattern at the root of the final term of D'_β. The result now follows by the induction hypothesis and Lemma 8.17 applied to D'_β. □

We can now prove Lemma 9.5:

Proof of Lemma 9.5. Let $s \twoheadrightarrow t$ be arbitrary. By compression and strong convergence, we may write $s \to^* t' \twoheadrightarrow^{\leq \omega} t$ such that all root-reductions occur in $s \to^* t'$. By repeated application of Lemma 9.7, there exists a hypercollapsing reduction starting from t'. In particular, t' reduces to a collapsing redex. Since $t' \twoheadrightarrow t$ contains no steps at the root, Lemma 9.8 yields that t reduces to a collapsing redex, proving that s is hypercollapsing. □

9.2 Confluence Modulo

We now prove confluence modulo identification of hypercollapsing subterms. Confluence modulo is defined as follows:

Definition 9.9. An iCRS is said to be *confluent modulo* an equivalence relation $\sim$ if $s \twoheadrightarrow s'$ and $t \twoheadrightarrow t'$ with $s \sim t$ implies existence of terms s'' and t'' such that $s' \twoheadrightarrow s''$ and $t' \twoheadrightarrow t''$ with $s'' \sim t''$ (see Figure 6).

We first show that identification of hypercollapsing subterms yields an equivalence relation. To this end we introduce some notation and show that hypercollapsingness is preserved under replacement of hypercollapsing subterms.

Notation 9.10. We write $s \sim_{hc} t$ if t can be obtained from s by replacing a number of hypercollapsing subterms of s by other hypercollapsing terms.

Proposition 9.11. *Let s and t be terms. If s is hypercollapsing and $s \sim_{hc} t$, then t is hypercollapsing.*

Proof. Let P be the set of positions of hypercollapsing subterms in s that are replaced to obtain t. By definition of s there exists a hypercollapsing reduction S starting from it. The redex patterns employed in the steps of S either occur completely outside or completely inside the reducts of the subterms at descendants of positions in P. This follows by orthogonality and the fact that the subterms at positions in P are hypercollapsing, i.e. each reduct reduces to a term with a collapsing redex at the root. It is irrelevant that any terms are substituted into the reducts of the subterms at descendants of positions in P along S by orthogonality and the fact that free variables cannot get bound when substituted in the reducts.

Omit from S all steps that occur inside the reducts of the subterms at descendants of positions in P to obtain a reduction S' of length α. By definition of S', together with orthogonality and fully-extendedness, there exists a reduction T of length α starting in t such that for all $\beta \leq \alpha$ we have that the redex pattern and position of the redex contracted in the βth step of both S' and T are identical. Hence, if S' is hypercollapsing then so T is and the result follows by Lemma 9.5. If S' is not hypercollapsing, then s reduces to a subterm at a position $p \in P$ and the same holds for T. As the subterm at position p in t is hypercollapsing, there exist a hypercollapsing reduction starting from it. Again, it is irrelevant that any terms are substituted in the subterm by orthogonality and the fact that free variables cannot get bound when substituted. Hence, T can be prolonged to obtain a hypercollapsing reduction and the result follows again by Lemma 9.5. □

We can now prove that $\sim_{hc}$ has the required properties:

Proposition 9.12. *The relation $\sim_{hc}$ is an equivalence relation, which is closed under substitution of terms for free variables.*

Proof. We have to prove that the relation is reflexive, symmetric, and transitive. Reflexivity and symmetry are immediate by definition. Transitivity follows by Proposition 9.11.

To see that relation is closed under substitution, consider a hypercollapsing term s and a term t that is a substitution instance of s. By definition of s there exists a hypercollapsing reduction S of length α starting from it. By orthogonality and the fact that no free variables are bound in the terms substituted in s, there exists a reduction T of length α starting for t such that for all $\beta \leq \alpha$ we have that the redex pattern and position of the redex contracted in βth step of both S and T are identical. Hence, since S is hypercollapsing, so is T and the result follows by Lemma 9.5. □

Introducing some further notation, we next show that we can accurately 'simulate' reductions in terms that are $\sim_{hc}$-related.

Notation 9.13. By $s \to^{\text{out}} t$ we denote a rewrite step that does not occur inside any hypercollapsing subterm of s.

Lemma 9.14. *Let $s \twoheadrightarrow t$ such that α steps along the reduction occur outside hypercollapsing subterms. If $s \sim_{hc} s'$, then there exists a reduction $s' \twoheadrightarrow^{\text{out}} t'$ of length α such that $t \sim_{hc} t'$. Moreover, for all $\beta \leq \alpha$ it holds that the redex pattern and position of the redex contracted in the βth step of $s' \twoheadrightarrow^{\text{out}} t'$ are identical to those of the βth step of $s \twoheadrightarrow t$ that occurs outside a hypercollapsing subterm.*

Proof. Let $s \twoheadrightarrow^{\gamma} t$ and $s \sim_{hc} s'$. We prove the result by ordinal induction on γ.

If $\gamma = 0$, then the result is immediate, as an empty reduction is by definition one that only contracts redexes outside hypercollapsing subterms.

If $\gamma = \delta + 1$, then assume $s \twoheadrightarrow^{\gamma} t = s \twoheadrightarrow^{\delta} s_\delta \to t$. By the induction hypothesis there exists a term s'_δ such that $s' \twoheadrightarrow^{out} s'_\delta$ and $s_\delta \sim_{hc} s'_\delta$. There are two possibilities for $s_\delta \to t$, depending on the contracted redex occurring either outside all hypercollapsing subterms or inside one of them:

- If the redex occurs outside all hypercollapsing subterms, then $s_\delta \sim_{hc} s'_\delta$ together with orthogonality and fully-extendedness implies that a redex employing the same rewrite rule as the redex contracted in $s_\delta \to t$ occurs at the same position in s'_δ. Moreover, the redex occurs outside all hypercollapsing subterms by Proposition 9.11. Hence, contracting the redex in s'_δ yields a step $s'_\delta \to^{\text{out}} t'$. That $t \sim_{hc} t'$ follows by $s_\delta \sim_{hc} s'_\delta$ and the fact that the same rewrite rule is employed in both $s_\delta \to t$ and $s'_\delta \to^{\text{out}} t'$: Clearly, t and t' are identical at all positions p that descend from positions not in hypercollapsing subterms of s_δ or s'_δ. If q is the position of a maximal hypercollapsing subterm of s_δ, then it is also the position of a maximal hypercollapsing subterm of s'_δ and vice versa, by Proposition 9.11. The descendants of q occur at identical positions in t and t' and are (not necessarily maximal) hypercollapsing subterms, since $s_\delta \sim_{hc} s'_\delta$ and since $\sim_{hc}$ is closed under substitution.

- If the redex occurs inside a hypercollapsing subterm, then $t \sim_{hc} s_\delta$. Hence, by transitivity of $\sim_{hc}$ we have $t \sim_{hc} s'_\delta$ and we can define $t' = s'_\delta$.

If γ is a limit ordinal, then the result is immediate by strong convergence and the induction hypothesis. □

Before proving the main theorem of this section, we show that reductions outside hypercollapsing subterms are confluent *modulo* $\sim_{hc}$. To this end we first prove a restricted variant of the Strip Lemma. It is of course well-known that the usual Strip Lemma for iTRSs fails for iλc, and hence we see that it must also fail for iCRSs.

Lemma 9.15 (Restricted Strip Lemma)**.** *If $S : s \twoheadrightarrow^{\text{out}} t$ and $T : s \to^{\text{out}} t'$, then S/T and T/S exist and end in the same term.*

Proof. Denote the length of S by α. We prove the Restricted Strip Lemma by ordinal induction on α. Note that, since T contracts a single redex u, we have that T/S is actually a complete development of the residuals of u in t. Obviously, if $\alpha = 0$, then the result follows trivially.

If α is a successor ordinal, then the result is immediate by Proposition 6.16 and the induction hypothesis.

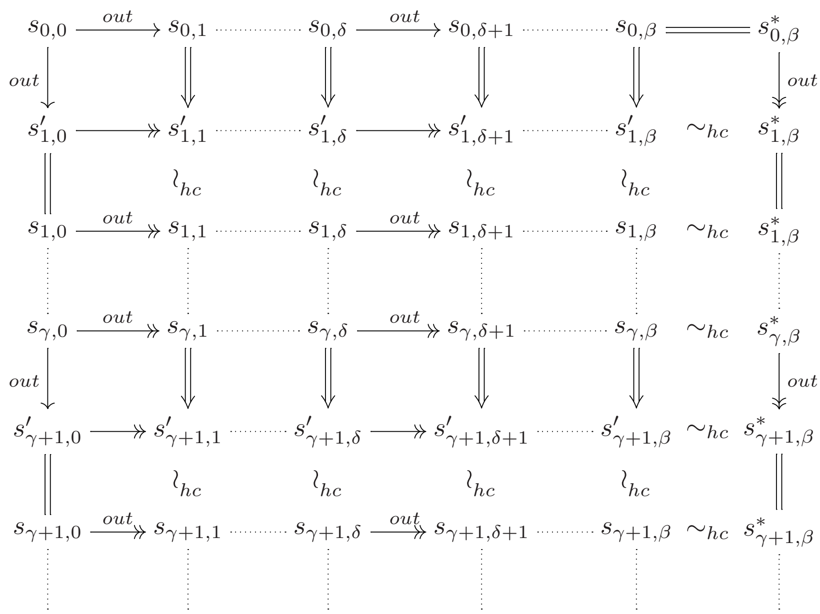
If α is a limit ordinal, then Theorem 7.2 and the induction hypothesis ensure that we only need to prove that T/S is strongly convergent. In other words, since T contracts a single redex u, we need to prove that u/S has a strongly convergent complete development. Assume the contrary. Observe that the rewrite rule employed in T must be collapsing, otherwise any development of u/S is strongly convergent. As contracting residuals of u cannot create further nestings of the residuals that are left, there exists a subterm of t with a hypercollapsing reduction starting from it (obtained by a development of u/S), say at position p. In fact, since no nestings are created, there must exist an infinite chain of residuals of u at p. By strong convergence and limit ordinal length of S, we can write $S = S_0; S_1$, where S_0 has successor ordinal length and where $S_1 : t^* \twoheadrightarrow^{\text{out}} t$ is a non-empty final segment of S performing no steps at prefix positions of p. Hence, S_0 has length strictly less than α and $t^*|_p \twoheadrightarrow^{\text{out}} t|_p$. As there is a hypercollapsing reduction starting from $t|_p$, it follows by Definition 9.2 that there is also a hypercollapsing reduction starting from $t^*|_p$. But then, by Lemma 9.5, we have that $t^*|_p$ is hypercollapsing, which implies that $t^*|_p \twoheadrightarrow^{\text{out}} t|_p$ is empty and that $t^*|_p = t|_p$. Thus, $t^*|_p$ contains a set of descendants of u having no complete development (giving rise to the hypercollapsing reduction starting from $t^*|_p = t|_p$), whence u/S_0 has no complete development. Since S_0 has length strictly less than α, this contradicts the induction hypothesis. Hence, T/S is strongly convergent. □

Lemma 9.16. *If $s \twoheadrightarrow^{\text{out}} t_0$ and $s \twoheadrightarrow^{\text{out}} t_1$, then there exist terms t_0^* and t_1^* such that $t_0 \twoheadrightarrow^{\text{out}} t_0^*$ and $t_1 \twoheadrightarrow^{\text{out}} t_1^*$ with $t_0^* \sim_{hc} t_1^*$.*

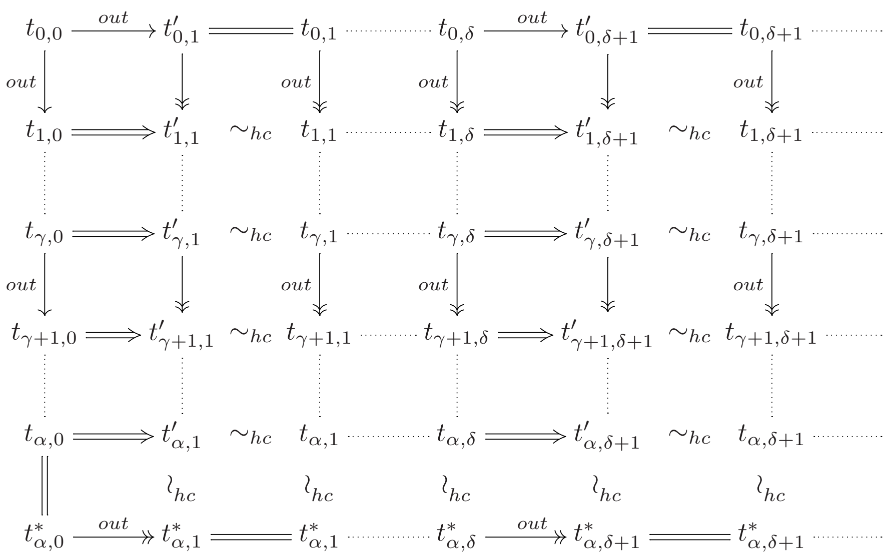
Proof. Let $S : s \twoheadrightarrow^{\text{out}} t_0$ and $T : s \twoheadrightarrow^{\text{out}} t_1$. By compression and Lemma 9.14 we may assume that both S and T have length at most ω. Suppose S has length $\alpha \leq \omega$ and T has length $\beta \leq \omega$. The proof proceeds in four steps: In the first step two 'tiling diagrams' are constructed, yielding respectively a reduction starting in t_0 and a reduction starting in t_1. In the second step a relation is established between the 'tiles' of the two diagrams. Employing the relation, it is shown in the third step that the two reductions obtained in the first step are strongly convergent. Finally, in the fourth step it is shown that the final terms of the two strongly convergent reductions are equivalent modulo $\sim_{hc}$.

Write $S : s_{0,0} \to^{\text{out}} s_{1,0} \to^{\text{out}} \cdots s_{\gamma,0} \to^{\text{out}} s_{\gamma+1,0} \to^{\text{out}} \cdots s_{\alpha,0}$ and $T : s_{0,0} \to^{\text{out}} s_{0,1} \to^{\text{out}} \cdots s_{0,\delta} \to^{\text{out}} s_{0,\delta+1} \to^{\text{out}} \cdots s_{0,\beta}$ and define $s'_{\gamma,0} = s_{\gamma,0}$ for all $\gamma \leq \alpha$. We inductively construct the 'tiling diagram' in Figure 7(a):

- the tiling of $s_{\gamma,0} \to^{\text{out}} s'_{\gamma+1,0}$ and $s_{\gamma,0} \twoheadrightarrow^{\text{out}} s_{\gamma,\beta}$ exists by Lemma 9.15;
- the reduction $s_{\gamma+1,0} \twoheadrightarrow^{\text{out}} s_{\gamma+1,\beta}$ and the equivalences $s_{\gamma+1,\delta} \sim_{hc} s'_{\gamma+1,\delta}$ for all $0 \leq \delta \leq \beta$ exist by Lemma 9.14 and the existence of $s'_{\gamma+1,0} \twoheadrightarrow s'_{\gamma+1,\beta}$;



(a)



(b)

Figure 7: The 'tiling diagrams' from the proof of Lemma 9.16

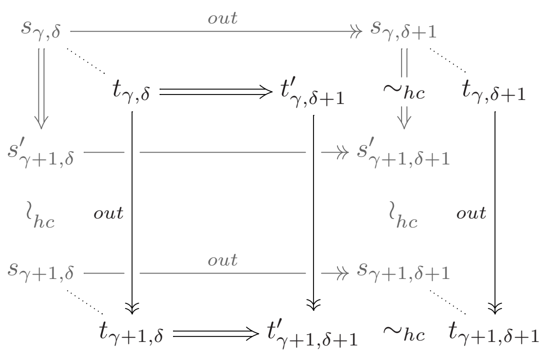


Figure 8: Superimposing the 'tiles' of the 'tiling diagrams' in Figure 7

- the reduction $s^*_{\gamma,\beta} \twoheadrightarrow^{\text{out}} s^*_{\gamma+1,\beta}$ and the equivalence $s'_{\gamma+1,\beta} \sim_{hc} s^*_{\gamma+1,\beta}$ exist by Lemma 9.14 and the existence of $s_{\gamma,\beta} \twoheadrightarrow^{\text{out}} s'_{\gamma+1,\beta}$;
- the equivalence $s^*_{\gamma+1,\beta} \sim_{hc} s_{\gamma+1,\beta}$ exists by transitivity of $\sim_{hc}$ and since $s_{\gamma+1,\beta} \sim_{hc} s'_{\gamma+1,\beta}$ and $s'_{\gamma+1,\beta} \sim_{hc} s^*_{\gamma+1,\beta}$.

As can be seen in Figure 7(a), the construction yields a reduction S^* starting in $t_1 = s^*_{0,\beta}$ such that all steps in the reduction occur outside hypercollapsing subterms. Note that the constructed diagram is not a tiling diagram in the strict sense of the word: No reduction occurs at the bottom and the diagram consists not only of reductions but also of equivalences modulo hypercollapsing subterms.

To obtain the second 'tiling diagram', depicted in Figure 7(b), we write $S : t_{0,0} \rightarrow^{\text{out}} t_{1,0} \rightarrow^{\text{out}} \cdots t_{\gamma,0} \rightarrow^{\text{out}} t_{\gamma+1,0} \rightarrow^{\text{out}} \cdots t_{\alpha,0}$ and $T : t_{0,0} \rightarrow^{\text{out}} t_{0,1} \rightarrow^{\text{out}} \cdots t_{0,\delta} \rightarrow^{\text{out}} t_{0,\delta+1} \rightarrow^{\text{out}} \cdots t_{0,\beta}$ and define $t'_{0,\delta} = t_{0,\delta}$ for all $\delta \leq \beta$. The diagram is constructed by vertically repeating the horizontal construction of Figure 7(a). The construction yields a reduction $T^* : t_0 = t^*_{\alpha,0} \twoheadrightarrow^{\text{out}} t^*_{\alpha,1} \twoheadrightarrow^{\text{out}} \cdots t^*_{\alpha,\delta} \twoheadrightarrow^{\text{out}} \cdots$.

Superimpose the tiles of the constructed 'tiling diagrams' as depicted in Figure 8, i.e. $s_{\gamma,\delta}$ and $t_{\gamma',\delta'}$ are superimposed if $\gamma = \gamma'$ and $\delta = \delta'$. Define $s_{0,\delta} = s'_{0,\delta}$, and $t_{\gamma,0} = t'_{\gamma,0}$ for all $\gamma \leq \alpha$ and $\delta \leq \beta$. By construction of the 'tiling diagrams', no term is superimposed on $s_{\gamma,\beta}$ with $\gamma \leq \alpha$ in case $\beta = \omega$ and similarly for $t_{\alpha,\delta}$ with $\delta < \beta$ in case $\alpha = \omega$.

We next prove for all superimposed terms $s_{\gamma,\delta}$ and $t_{\gamma,\delta}$ that $s_{\gamma,\delta} \sim_{hc} s'_{\gamma,\delta} \sim_{hc} t_{\gamma,\delta} \sim_{hc} t'_{\gamma,\delta}$. The proof is by induction on γ and δ. Induction is allowed because $s_{\gamma,\delta}$ and $t_{\gamma,\delta}$ exist for all $\gamma < \alpha$ and $\delta < \beta$:

- In case either $\gamma = 0$ *or* $\delta = 0$, we defined $s_{\gamma,\delta} = s'_{\gamma,\delta} = t_{\gamma,\delta} = t'_{\gamma,\delta}$. Hence, since $\sim_{hc}$ is an equivalence relation, $s_{\gamma,\delta} \sim_{hc} s'_{\gamma,\delta} \sim_{hc} t_{\gamma,\delta} \sim_{hc} t'_{\gamma,\delta}$.
- In case of $\gamma = \gamma' + 1$ *and* $\delta = \delta' + 1$, we have by definition of the 'tiling diagrams' that $s_{\gamma,\delta} \sim_{hc} s'_{\gamma,\delta}$ and $t_{\gamma,\delta} \sim_{hc} t'_{\gamma,\delta}$. Hence, by transitivity of

$\sim_{hc}$, we obtain the desired result if we can establish $s_{\gamma,\delta} \sim_{hc} t'_{\gamma,\delta}$.

By Lemmas 9.15 and 9.14, as employed in the construction of the 'tiling diagrams', $s_{\gamma,\delta'} \twoheadrightarrow^{\text{out}} s_{\gamma,\delta}$ is essentially a development of residuals of the redex u contracted in $s_{0,\delta'} \rightarrow^{\text{out}} s_{0,\delta}$ such that no residuals of u in $s_{\gamma,\delta}$ remain *outside* hypercollapsing subterms. Since we have by the induction hypothesis that $s_{\gamma,\delta'} \sim_{hc} t_{\gamma,\delta'}$ and since every step in $s_{\gamma,\delta'} \twoheadrightarrow^{\text{out}} s_{\gamma,\delta}$ occurs outside hypercollapsing subterms, it follows by orthogonality and fully-extendedness that there exists a reduction $t_{\gamma,\delta'} \twoheadrightarrow t''_{\gamma,\delta}$ such that $s_{\gamma,\delta} \sim_{hc} t''_{\gamma,\delta}$. Since $s_{\gamma,\delta'} \twoheadrightarrow^{\text{out}} s_{\gamma,\delta}$ is essentially a development of residuals of u, it follows that $t_{\gamma,\delta'} \twoheadrightarrow t''_{\gamma,\delta}$ is development of residuals of u, i.e. of the redex contracted in $t_{0,\delta'} \rightarrow t'_{0,\delta}$. Moreover, it follows that all residuals of u left in $t''_{\gamma,\delta}$ occur inside hypercollapsing subterms. Hence, since we have by Lemma 6.14 that $t''_{\gamma,\delta} \twoheadrightarrow t'_{\gamma,\delta}$, we also have that $t''_{\gamma,\delta} \sim_{hc} t'_{\gamma,\delta}$. But then, by transitivity of $\sim_{hc}$ it follows that $s_{\gamma,\delta} \sim_{hc} t'_{\gamma,\delta}$, as required.

Employing that $s_{\gamma,\delta} \sim_{hc} s'_{\gamma,\delta} \sim_{hc} t_{\gamma,\delta} \sim_{hc} t'_{\gamma,\delta}$ holds for all superimposed $s_{\gamma,\delta}$ and $t_{\gamma,\delta}$, we next prove that the reduction $S^* : s^*_{0,\beta} \twoheadrightarrow^{\text{out}} s^*_{1,\beta} \twoheadrightarrow^{\text{out}} \cdots s^*_{\gamma,\beta} \twoheadrightarrow^{\text{out}} \cdots$ in Figure 7(a) is strongly convergent. The proof is by contradiction. Thus, suppose S^* is not strongly convergent. There now exists a position p of minimal depth d such that an infinite number of reductions occur at p. As each step in S^* occurs outside hypercollapsing subterms, it follows by minimality of d that from some γ onwards no redexes are contracted above p and that all redexes contracted at p are non-collapsing. Moreover, by strong convergence of $s_{\gamma,0} \twoheadrightarrow s_{\gamma,\beta}$, there exist a δ such that all steps in $s_{\gamma,\delta} \twoheadrightarrow s_{\gamma,\beta}$ also occur below d.

Suppose for some minimal $\kappa \geq \gamma$ that a redex is contracted at some position $q < p$ in either $s_{\kappa,\delta} \Rightarrow s'_{\kappa+1,\delta}$ or $s_{\kappa,\delta} \twoheadrightarrow s_{\kappa,\beta}$. By dependence of the depth of the steps in $s_{\kappa,\delta} \twoheadrightarrow s_{\kappa,\beta}$ on the depth of the steps in $s_{\lambda,\delta} \twoheadrightarrow s_{\lambda,\beta}$ for all $\gamma \leq \lambda < \kappa$, it follows by minimality of κ that the reduction must be $s_{\kappa,\delta} \Rightarrow s'_{\kappa+1,\delta}$. This implies that a redex is also contracted at position q in $s_{\kappa,\beta} \Rightarrow s'_{\kappa+1,\beta}$. Since the redex is by definition not contracted in $s^*_{\kappa,\beta} \twoheadrightarrow^{\text{out}} s^*_{\kappa+1,\beta}$, it follows that the subterm at position q in $s^*_{\kappa,\beta}$ is hypercollapsing. However, as $q < p$, this implies that the infinite number of redexes contracted at position p cannot occur, as redexes in S^* are contracted outside hypercollapsing subterms. Hence, for all $\kappa \geq \gamma$ we have that no reduction $s_{\kappa,\delta} \Rightarrow s'_{\kappa+1,\delta}$ or $s_{\kappa,\delta} \twoheadrightarrow s_{\kappa,\beta}$ contracts a redex at strict prefix position of p.

Since all steps in $s_{\gamma,\delta} \twoheadrightarrow s_{\gamma,\beta}$ occur below d, the previous implies that if a redex is contracted at position p in some $s^*_{\kappa,\beta} \twoheadrightarrow^{\text{out}} s^*_{\kappa+1,\beta}$ for minimal $\kappa \geq \gamma$, then a redex is also contracted at position p in $s_{\kappa,\delta} \Rightarrow s'_{\kappa+1,\delta}$. Since the contracted redex is non-collapsing, it follows that the function symbol that occurs at position p in both $s^*_{\kappa+1,\beta}$ and $s'_{\kappa+1,\delta}$ is the root symbol of the next redex contracted at position p. Hence, $s_{\gamma,\delta} \Rightarrow s'_{\gamma+1,\delta} \sim_{hc} s_{\gamma+1,\delta} \Rightarrow s'_{\gamma+2,\delta} \sim_{hc} s_{\gamma+2,\delta} \Rightarrow \cdots$ contains an infinite number of steps at position p without any interleaving of collapsing steps at that position. However, as redexes contracted at position p cannot occur inside hypercollapsing subterms by definition of S^*, we have that $t_{0,\delta} \twoheadrightarrow t_{\alpha,\delta}$ also contracts an infinite number of redexes at position

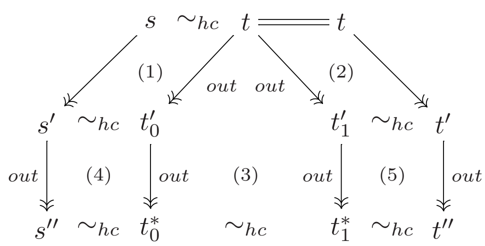
p, which is impossible by strong convergence of this reduction, contradiction. Hence, S^* is strongly convergent.

By a similar argument as above it follows that the reduction $T^* : t^*_{\alpha,0} \twoheadrightarrow^{\text{out}} t^*_{\alpha,1} \twoheadrightarrow^{\text{out}} \cdots t^*_{\alpha,\delta} \twoheadrightarrow^{\text{out}} \cdots$ is strongly convergent. Hence, since $s_{\gamma,\delta} \sim_{hc} s'_{\gamma,\delta} \sim_{hc} t_{\gamma,\delta} \sim_{hc} t'_{\gamma,\delta}$ for all γ and δ in both 'tiling diagrams', the desired result follows by strong convergence. □

We can now — finally — prove the main result of the section: confluence modulo $\sim_{hc}$.

Theorem 9.17. *Fully-extended, orthogonal iCRSs are confluent modulo $\sim_{hc}$.*

Proof. Let $s \sim_{hc} t$, and assume that $s \twoheadrightarrow s'$ and $t \twoheadrightarrow t'$. Consider the following diagram:



In the diagram, (1) and (2) exist by Lemma 9.14 and (3) exists by Lemma 9.16. Moreover, (4) and (5) also exist by Lemma 9.14. The result now follows by the diagram and transitivity of $\sim_{hc}$. □

9.3 Almost Non-Collapsingness

We would like a characterisation of confluence that appeals only to the syntax of iCRSs without any need to consider equality modulo some relation. The first correct, fundamental confluence result for iTRSs [18] stated that an orthogonal iTRSs is confluent iff it has the property of being 'almost non-collapsing': There is at most one rule that is collapsing and the variable at the root of the right-hand side of that rule is the only variable occurring in the left-hand side of that rule.

Unfortunately, this concept does not carry over trivially to iCRSs, when replacing the variables from iTRSs by meta-variables:

Example 9.18. Consider the following rewrite rule, which is almost non-collapsing in the above sense:

$$f([x]Z(x)) \to Z(f([x]Z(x)) .$$

The term $f([x]f([y]x))$ gives rise to the finite reduction

$$f([x]f([y]x)) \to f([x]x) ,$$

the final term of which reduces only to itself. However, the following reduction of length ω also exists:

$$f([x]f([y]x)) \to f([y]f([x]f([y]x))) \to f([y]f([y]f([x]f([y]x)))) \to \cdots s\,,$$

where s is solution of the recursive equation $s = f([y]s)$, which is again a term which only reduces to itself. Hence, $f([x]f([y]x))$ reduces to two different terms that only reduce to themselves. In other words, the considered 'almost non-collapsing' rewrite rule defines a non-confluent iCRS.

We currently do not know how to give a precise characterisation of the class of confluent iCRSs. From the above example, it is clear that almost non-collapsingness alone does not suffice. It is plausible that the criterion for confluence will be undecidable, even for the class of iCRSs containing only a finite number of rules, all of which have finite right-hand-sides. The above example bears witness of this: It crucially depends on the redex $f([x]x)$ being reachable from itself and reachability is of course in general undecidable.

10 Normal Forms and Normalisation

In this section we consider normal forms of iCRSs:

Definition 10.1. A term in an iCRS is a *normal form* if no redexes occur in the term.

In Section 10.1 we consider several properties of normal forms. Thereafter, in Sections 10.2, 10.3, and 10.4 we resume to consider fully-extended, orthogonal iCRSs when we apply the technique of *essentiality* in the context in which it was originally conceived: that of normalising reduction strategies. The considered reduction strategies are all 'fair' reductions. We consider outermost-fair reductions, fair reductions, and needed-fair reductions. These reductions all satisfy the following definition:

Definition 10.2. Let $\mathcal{P}$ be a predicate. A *$\mathcal{P}$-fair reduction* is a weakly continuous reduction $(s_\beta)_{\beta<\alpha}$ where for every $\beta < \alpha$ and redex u in s_β satisfying $\mathcal{P}$ there exists a $\beta \leq \gamma < \beta + \omega$ such that either

1. $s_\gamma \to s_{\gamma+1}$ contracts a residual of u satisfying $\mathcal{P}$, or
2. no residual of u in s_γ satisfies $\mathcal{P}$.

Thus, each redex that satisfies $\mathcal{P}$ is either reduced after a finite number of steps or it no longer satisfies $\mathcal{P}$ after a finite number of steps.

Example 10.3. Given that a redex at a position p is called outermost if no redex occurs at a strict prefix position of p, we can define *outermost-fair* reductions by defining a predicate for redexes that is true in case the redex is outermost and false otherwise. Consider the following two rewrite rules:

$$f(Z) \to g(Z)$$
$$a \to g(a)$$

Next, consider the reduction

$$f(a) \to f(g(a)) \to \cdots \to f(g^n(a)) \to g^{n+1}(a) \to \cdots g^\omega$$

and

$$f(a) \to f(g(a)) \to \cdots \to f(g^n(a)) \to \cdots \to f(g^\omega) \to g^\omega .$$

The first of these reductions is outermost-fair, as a residual of each redex present in each term is reduced after a finite number of steps. The second reduction is *not* outermost-fair, as a residual of a redex that occurs at the root of $f(a)$ is only contracted after ω steps and as a redex occurring at the root of a term is always outermost.

10.1 Normal Form Properties

The following properties relate the normal forms of an iCRS and its rewrite relation. The properties extend their usual finitary counterparts to infinitary rewriting. Ample motivation for the formulation of the properties can be found in [18].

Definition 10.4. Define the following:

- An iCRS has the *normal form property (NF)* if $s \; (\twoheadleftarrow \cdot \twoheadrightarrow)^* \; t$ with t a normal form implies $s \twoheadrightarrow t$, where $(\twoheadleftarrow \cdot \twoheadrightarrow)^*$ denotes the symmetric, transitive, reflexive closure of $\twoheadrightarrow$.
- An iCRS has the *unique normal form property (UN)* if $s \; (\twoheadleftarrow \cdot \twoheadrightarrow)^* \; t$ with s and t normal forms implies $s = t$.
- An iCRS has the *unique normal form property with respect to reduction (UN$^\rightarrow$)* if $t \twoheadleftarrow s \twoheadrightarrow t'$ with t and t' normal forms implies $t = t'$.

By the definitions we immediately have:

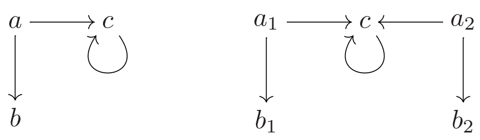
Proposition 10.5. *It holds that NF implies UN and that UN implies UN$^\rightarrow$.*

The reverse implications of those above do not hold. This can be witnessed by the rewrite systems depicted in Figure 9.

In Figure 9(a) a counterexample occurs refuting that UN implies NF: Since b is the only normal form, next to all variables, UN is immediate. However, NF does not hold, as there is no reduction $c \twoheadrightarrow b$. The rewrite system in Figure 9(b) refutes that UN$^\rightarrow$ implies UN: Since b_1 is the only normal form of a_1 with respect to reduction and since b_2 the only normal form of a_2, UN$^\rightarrow$ is immediate. However, UN does not hold, as we have $b_1 \; (\twoheadleftarrow \cdot \twoheadrightarrow)^* \; b_2$, while $b_1 \neq b_2$.

The following lemma relates confluence modulo hypercollapsing subterms with the three properties introduced above.

Lemma 10.6. *If an iCRS is confluent modulo hypercollapsing subterms, then NF, UN, and UN$^\rightarrow$ hold.*



(a) (b)

Figure 9: Counterexamples to the reverse of Proposition 10.5

Proof. Let $s\ (\twoheadleftarrow\cdot\twoheadrightarrow)^*\ t$ with t a normal form. By induction on the number of changes in the direction of the rewrite relation in $s\ (\twoheadleftarrow\cdot\twoheadrightarrow)^*\ t$ and confluence modulo hypercollapsing subterms it follows that s reduces to a term t' such that $t \sim_{hc} t'$. As no hypercollapsing subterms occur in normal forms, we have $t = t'$. Hence, NF holds and UN and $\text{UN}^{\rightarrow}$ follow by Proposition 10.5. ☐

In the above proof confluence is easily substituted for confluence modulo hypercollapsing subterms, yielding the traditional result from finitary rewriting stating that confluence implies NF, UN, and $\text{UN}^{\rightarrow}$. Moreover, as fully-extended, orthogonal iCRSs are confluent modulo hypercollapsing subterms by Theorem 9.17, the above lemma also gives an affirmative answer to the conjecture posed in [23] stating that fully-extended, orthogonal iCRSs satisfy $\text{UN}^{\rightarrow}$.

It is not the case that NF implies confluence modulo hypercollapsing subterms. To see this, consider the following four rewrite rules:

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

No term in which a redex occurs has a normal form. Hence, NF is immediate. However, confluence modulo hypercollapsing subterms does not hold, as a reduces to $f(b)$ and $g(c)$, both of which only reduce to themselves, and as $f(b) \not\sim_{hc} g(c)$.

10.2 Outermost-Fair Reductions

We desire a method to obtain normal forms in infinitary rewriting. The standard way of obtaining normal forms in (finitary) higher-order rewriting is by using an *outermost-fair* strategy [30, 39, 38].

Definition 10.7. Let s be a term. A redex at a position p in s is *outermost* if no redex in s occurs at a strict prefix position of p. An *outermost-fair reduction* is a $\mathcal{P}$-fair reduction such that $\mathcal{P}$ is true for a redex iff the redex is outermost.

Hence, a reduction is outermost-fair if, after a finite number of steps, every outermost redex is either reduced or is not outermost anymore.

Lemma 10.8. *Let s be a term and T an outermost-fair reduction of length at least ω starting in s. If there is a reduction $s \twoheadrightarrow t$ with t a normal form, then T is strongly convergent of length ω with limit t.*

Proof. By compression, we may assume that $s \twoheadrightarrow t$ has length at most ω. Moreover, by strong convergence we may write $s \twoheadrightarrow t$ as:

$$s \to^* s_1 \to^* \cdots \to^* s_d \to^* \cdots t\,,$$

where all steps in $s_d \twoheadrightarrow t$ occur below depth d. For each depth $d > 0$ and $D_d : s \to^* s_d$, we have by definition that s_d mirrors t in P_d, where P_d is the set of positions in t above depth d. Moreover, no redexes occur in s_d at positions in P_d and, as $s \to^* s_d$ is finite, we can view D_d to be a finite sequence of complete developments, where each development consists of a single step. Hence, $\mu_{P_d}(D_d)$ as defined in Section 8 exists.

Let the depth $d > 0$ be arbitrary and denote the first ω steps of T by T_ω. Iteratively, take the emaciated projection of D_d over this initial sequence. By Lemmas 8.15 and 8.16 and well-foundedness of $\prec$, only a finite number of steps of T_ω are essential for P_d. Following the finite number of essential steps, there are two possibilities for the emaciated projection of D_d: Either all the developments in the projection are empty, or not.

- In case all the developments are empty it follows by Lemma 8.15 that all remaining terms along T_ω mirror s_d and t in P_d and that no redexes are contracted above depth d.

- In case not all the developments are empty, it follows by Lemma 8.15 that exists a fixed set of essential positions P such that all the remaining terms along T_ω mirror each other in P. Moreover, the lemma together with non-emptiness implies that a redex u occurs at a fixed position in P. Since the depth of u is finite, only a finite number of redexes can be created above u in the remaining part of T_ω. These redexes cannot be contracted or cease to exist by orthogonality and since all further contracted redexes occur at positions not in P, again by Lemma 8.15. Hence, after a finite number of further steps a redex must be created that is outermost for the remainder of T_ω, contradicting outermost-fairness. Thus, the emaciated projection of D_d must become empty after a finite number of steps.

As the previous holds for all depths $d > 0$ we have that T_ω is strongly convergent with limit t. Hence, $T = T_\omega$ and the result follows. □

We thus obtain a strong result concerning normalisation of iCRSs:

Theorem 10.9. *If s can be reduced to normal form by a strongly convergent reduction, then it also reduces to a normal form by any outermost-fair reduction. Any such reduction is strongly convergent and of length at most ω.*

Proof. If T is a finite outermost-fair reduction starting in s and T reaches a normal form, then we are done. If T is finite but has not reached a normal form, then there is at least one outermost redex in the final term of T, and we may thus extend it. Hence, we only need to prove that if S is infinite, that S is also strongly convergent of length ω and reaches a normal form. This is the contents of Lemma 10.8. □

10.3 Fair Reductions

Contrary to the predicate considered in the previous section, which is only satisfied under certain conditions, this section considers a predicate that will always be satisfied.

Definition 10.10. A *fair reduction* is a $\mathcal{P}$-fair reduction such that $\mathcal{P}$ is always true.

As the predicate is always true, the second clause of Definition 10.2 cannot occur in a fair reduction unless the first clause applies earlier on in the considered reduction.

We have the following:

Theorem 10.11. *If s can be reduced to a normal form by a strongly convergent reduction, then it also reduces to a normal form by any fair reduction. Any such reduction is strongly convergent and of length at most ω.*

Proof. Since a fair reduction is in particular fair with respect to outermost redexes, the result follows by Theorem 10.9. □

Hence, by the proof it follows that each fair reduction is an outermost-fair reduction. This implies that the predicate that is always true strengthens the predicate used for outermost-fair reductions, leading to a weaker result than the one obtained in the previous section.

10.4 Needed-Fair Reductions

In this section we show that needed-fair reductions are normalising.

Definition 10.12. Let s be a term. A redex u in s is *needed* if along every strongly convergent reduction from s to a normal form some residual of u is contracted. A *needed-fair reduction* is a $\mathcal{P}$-fair reduction such that $\mathcal{P}$ is true for a redex iff the redex is needed.

By definition of neededness, the second clause of Definition 10.2 cannot occur in a needed-fair reduction unless the first clause applies earlier on in the considered reduction.

Example 10.13. Consider the rewrite rules and reductions in Example 10.3. The first of the reductions is needed-fair, as each redex along the reduction is reduced after a finite number of steps. The second reduction is not needed-fair, as the redex at the root of $f(a)$ is only reduced after ω steps and as the redex at the root of a term is by definition needed.

To prove the normalisation of needed-fair reductions, we establish a relation between essential redexes and needed redexes. To this end we first relate the essential redexes along different complete developments of the same set of redexes:

Lemma 10.14. *Let s and t be terms, $\mathcal{U}$ a set of redexes of s such that $s \Rightarrow^{\mathcal{U}} t$, and P a prefix of t. If $s \Rightarrow^{\mathcal{V}_1} t' \Rightarrow^{\mathcal{V}_2} t$ with $\mathcal{V}_1 \subseteq \mathcal{U}$ and $\mathcal{V}_2 = \mathcal{U}/(s \Rightarrow t')$, then the set of positions essential for P in s is identical along $s \Rightarrow t$ and $s \Rightarrow t' \Rightarrow t$.*

Proof. By Lemma 6.11, $\mathcal{U}$ satisfies the finite jumps property. Hence, since $s \Rightarrow t$ and $s \Rightarrow s' \Rightarrow t$ are both complete developments of $\mathcal{U}$, we have by Theorem 6.10 that the set of descendants in t of a position in s is identical along both developments. Hence, by Proposition 8.8 it follows for any position in s with a descendant in P that the position is essential irrespective of the development being either $s \Rightarrow t$ or $s \Rightarrow s' \Rightarrow t$, where the proposition is applied twice in case of the latter development. This leaves to prove that the same holds for positions in redex patterns of redexes in $\mathcal{U}$.

Consider a fresh unary function symbol k and replace each subterm s' of s with a redex from $\mathcal{U}$ occurring at the root by $k(s')$. This yields a term s^k. Since the unary function symbol k does not occur in any of the rewrite rules of the assumed iCRS, it is easily shown for each (not necessarily complete) development starting in s that there exists a corresponding development starting in s^k, where the set of redexes is adapted appropriately and such that removal of all function symbols k yields the original development. Hence, the completeness of a development starting in s implies the completeness of the corresponding development starting in s^k.

Suppose that $s^k \Rightarrow t^k$ is the complete development that corresponds to $s \Rightarrow t$. Define the prefix P^k of t^k in such a way that the removal of the function symbols k from t^k and the corresponding elements from the positions in P^k yields P and such that for any position $p \in P^k$ that is not the prefix of any another position in P^k it holds that $root(t^k|_p) \neq k$. By definition of P^k and the definition of essentiality we immediately have that a redex in $\mathcal{U}$ is essential if the function symbol k directly preceding it in s^k is. Hence, by looking at the function symbols k occurring in s^k, the result now follows for the positions in the redex patterns of the redexes in $\mathcal{U}$ in similar fashion as for all other positions. □

By the previous lemma we also have the following:

Corollary 10.15. *Let s and t be terms, $\mathcal{U}$ a set of redexes of s such that $s \Rightarrow^{\mathcal{U}} t$, and P a prefix of t. If it holds that:*

- *$s \Rightarrow^{\mathcal{V}_1} s' \Rightarrow^{\mathcal{V}_2} t$ with $\mathcal{V}_1 \subseteq \mathcal{U}$ and $\mathcal{V}_2 = \mathcal{U}/(s \Rightarrow s')$, and*
- *$s \Rightarrow^{\mathcal{V}'_1} t' \Rightarrow^{\mathcal{V}'_2} t$ with $\mathcal{V}'_1 \subseteq \mathcal{U}$ and $\mathcal{V}'_2 = \mathcal{U}/(s \Rightarrow t')$,*

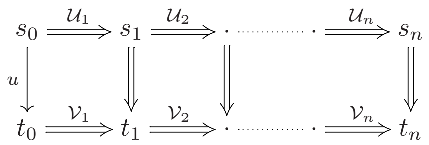
then the set of positions essential for P in s is identical along both $s \Rightarrow s' \Rightarrow t$ and $s \Rightarrow t' \Rightarrow t$.

We next show that each essential redex has an essential residual as long as it is not contracted and that inessential redexes only have inessential residuals. Moreover, we show that the same holds in case emaciated projections are considered.

Lemma 10.16. *Let $D : s_0 \Rightarrow^{\mathcal{U}_1} s_1 \Rightarrow^{\mathcal{U}_2} \cdots \Rightarrow^{\mathcal{U}_n} s_n$, with $\mathcal{U}_i$ finite for all $1 \leq i \leq n$, and let P be a prefix of s_n. If $s_0 \to t_0$ contracts a redex u such that no redex in u/D occurs at a position in P, then for every redex v in s_0:*

- *if v is essential, then it is either u or there exists a residual in t_0 that is essential, and*
- *if v is inessential, then all residuals in t_0 are inessential.*

Proof. Consider the following diagram:



where the reduction at the bottom is D/u. Since no redex in u/D occurs at a position in P, we have that t_n mirrors s_n in P. Hence, we can consider the redexes in t_0 essential for P. By repeated application of the Corollary 10.15 to the tiles of the diagram and by Proposition 8.8, it follows for all $1 \leq i \leq n$ that any redex in $\mathcal{U}_i$ that is not a residual of u has an essential residual in $\mathcal{V}_i$ in case it is essential and has no such residual in case it is inessential. Hence, by definition of essential redexes, every residual in t_0 of a redex in s_0 satisfies the required properties. □

Lemma 10.17. *Let $D : s_0 \Rightarrow^{\mathcal{U}_1} s_1 \Rightarrow^{\mathcal{U}_2} \cdots \Rightarrow^{\mathcal{U}_n} s_n$, with $\mathcal{U}_i$ finite for all $1 \leq i \leq n$, and let P be a prefix of s_n. If $s_0 \to t_0$ contracts a redex u such that no redex in u/D occurs at a position in P, then for every redex v in s_0:*

- *if v is essential, then it is either u or there exists an essential redex in the first term of $D /\!\!/ u$ with the redex pattern and position identical to that of a residual of v in t_0, and*
- *if v is inessential, then all redexes in the first term of $D /\!\!/ u$ with the redex pattern and position identical to that of a residual of v in t_0 are inessential.*

Proof. Immediate by the previous lemma and the definition of $D /\!\!/ u$. □

We are now in a position to relate essential redexes and needed redexes.

Lemma 10.18. *Let $s_0 \to^* s_1 \to^* \cdots \to^* s_d \to^* \cdots t$ be a reduction of length at most ω with t a normal form such that, for all $d \in \mathbb{N}$, the steps in $s_d \twoheadrightarrow t$ occur below depth d and P_d is the set of positions in s_d above depth d. If a redex in s is essential for some P_d in s with $d \in \mathbb{N}$, then it is needed.*

Proof. Let u be a redex in s that is essential for some prefix P_d. Moreover, suppose u is not needed. By definition of neededness there exists a reduction T to normal form not contracting any residual of u. We show by ordinal induction that a residual of u occurs in every term t_α along T. To facilitate the induction we also show for all $\beta \leq \alpha$ that each t_β reduces to a term that mirrors s_d in P_d by a finite sequence of complete developments D_β such that $\mu_{P_d}(D_\alpha) \preceq \mu_{P_d}(D_\beta)$ and such that s_α mirrors s_β in the positions of s_β essential for P_d in case $\mu_{P_d}(D_\alpha) = \mu_{P_d}(D_\beta)$.

For $t_0 = s_0$ the result is immediate by assumption. Since $s_0 \to^* s_d$ consists of a finite number of steps it is also a finite sequence of complete developments, where each development consists of a single step.

For $t_{\alpha+1}$, discriminate between the contracted redex, say v, being either essential or inessential. If v is essential, the result follows by Lemmas 8.16 and 10.17 and the induction hypothesis. If v is inessential, the result follows by Lemmas 8.15 and 10.17 and the induction hypothesis. In both cases the lemmas may be applied since v is not a residual of u by assumption of T and since all redexes occur below depth d in s_d and, hence, the same holds for the final term of D_α.

For s_α with α a limit ordinal, it follows by the well-foundedness of $\prec$ that there exist a $\beta < \alpha$ such that for every $\beta < \gamma < \alpha$ we have $\mu_{P_d}(D_\gamma) = \mu_{P_d}(D_\beta)$. Hence, by the induction hypothesis it follows for all $\beta < \gamma < \alpha$ that s_γ mirrors s_β in the positions of s_β essential for P_d and, by strong convergence, s_α mirrors s_β in these positions. By Lemma 8.13 there exists a finite sequence of complete developments D'_β starting in s_β contracting only essential redexes such that the final term of D'_β mirrors s_d in P_d. The result now follows by the induction hypothesis and Lemma 8.17 applied to D'_β.

Hence, a residual of u occurs in the final term of T, contradicting the fact that the final term is in normal form. Concluding, we have that every redex that is essential for some P_d is needed. □

Lemma 10.19. *Let $s_0 \to^* s_1 \to^* \cdots \to^* s_d \to^* \cdots t$ be a reduction of length at most ω with t a normal form such that, for all $d \in \mathbb{N}$, the steps in $s_d \twoheadrightarrow t$ occur below depth d and P_d is the set of positions in s_d above depth d. If a redex in s is needed, then there exists a $d \in \mathbb{N}$ such that the redex is essential for P_d in s.*

Proof. Consider a reduction S that contracts for increasingly larger $d \in \mathbb{N}$ all redexes essential for P_d until such redexes no longer occur. By Lemmas 8.15 and 8.16 and the well-foundedness of $\prec$ it follows that S is a strongly convergent reduction of length at most ω to the normal form t. Hence, given the existence of S and Lemma 10.17, a redex can only be needed in s if there exists some $d \in \mathbb{N}$ such that the redex is essential for P_d in s. □

Lemma 10.20. *Let $s_0 \to^* s_1 \to^* \cdots \to^* s_d \to^* \cdots t$ be a reduction of length at most ω with t a normal form such that, for all $d \in \mathbb{N}$, the steps in $s_d \twoheadrightarrow t$ occur below depth d and P_d is the set of positions in s_d above depth d. A redex in s is needed iff there exists a $d \in \mathbb{N}$ such that the redex is essential for P_d in s.*

Proof. Immediate by Lemmas 10.18 and 10.19. □

Finally, we can prove the main theorem of this section:

Theorem 10.21. *If s can be reduced to a normal form by a strongly convergent reduction, then it also reduces to a normal form by any needed-fair reduction. Any such reduction is strongly convergent and of length at most ω.*

Proof. By compression, we may assume we have a reduction $s \twoheadrightarrow t$ to normal form of length at most length ω. Write the reduction as

$$s = s_0 \rightarrow^* s_1 \rightarrow^* \cdots \rightarrow^* s_d \rightarrow^* \cdots t$$

with all steps in $s_d \twoheadrightarrow t$ occurring below depth d and denote by P_d the set of positions in s_d above depth d. Iteratively, consider the emaciated projection of a reduction $s \rightarrow^* s_d$ with respect to the prefix P_d. All redexes occurring at positions essential for P_d in the terms along the needed-fair reduction are essential, as no redexes occur in s_d at positions in P_d. Hence, it follows by Lemma 10.20 and the needed-fair condition that an essential redex is contracted after a finite number of steps as long as there are any essential redexes left (by Lemma 8.15 a prefix stays untouched as long as only inessential redexes are contracted). As $\prec$ is well-founded, we have by Lemmas 8.15 and 8.16 that the needed-fair reduction reduces s to a term that mirrors s_d in P_d in a finite number of steps. Since the previous holds for any s_d and P_d, it follows that any needed-fair reduction is strongly convergent and of length at most ω. □

11 Conclusion and Suggestions for Future Work

This paper has introduced and developed the theory of infinitary Combinatory Reduction Systems, thus providing the first true extension of infinitary rewriting to the higher-order setting. We have proven a number of results showing that many of the positive results from the ordinary (non-infinitary) setting can be lifted to infinitary systems; particularly useful results concern confluence and normalisation.

While the results of this paper generalise most known results in the field of infinitary rewriting, a number of important questions remain. We urge our readers to treat these at their leisure; an unprioritised list of open questions remaining is the following.

- Does there exist a notion of meaningless terms [19] that allows for the construction of Böhm-like trees?
- Can the treatment of iCRSs in this paper be extended to the other formats of higher-order rewriting? The fact that CRSs have a clean separation of abstractions (in terms and rewrite rules) and substitutions which is not present in some of the other forms of higher-order rewriting [40] may constitute a stumbling block in this respect.
- Does our proof of Theorem 5.2 construct compressed reductions Lévy-equivalent to the original ones. If not, is it possible to do so?

- Can a large subclass of higher-order iCRSs be identified for which a generalisation of the first-order result that almost-non-collapsing systems are confluent holds? As we show in Section 9.3, a generalisation is not likely to be easy to come by.

- The current proof of confluence modulo hypercollapsing subterms requires both fully-extendedness and orthogonality. Is it possible to drop the the requirement of fully-extendedness or replace orthogonality by weak orthogonality? Dropping fully-extendedness highly likely requires completely different proof techniques, since the assumption that reductions can be compressed to length at most ω is firmly embedded in the current essentiality approach. It ensures that the employed measure is well-founded.

- Apart from the obvious modelling of infinitary formulae and equations using terms and rules, do our results have any impact on infinitary (higher-order) logic with quantifiers [26]?

We hope that our readers will endeavour to answer the above and that the array of positive theorems in this paper will lead to further use of infinitary rewriting in the modelling of lazy data structures and lazy (declarative) languages.

References

[1] P. Aczel. A general Church-Rosser theorem. Technical report, University of Manchester, July 1978.

[2] A. Arnold and M. Nivat. The metric space of infinite trees. Algebraic and topological properties. *Fundamenta Informaticae*, 3(4):445–476, 1980.

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

[4] H. P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics*. Elsevier Science, revised edition, 1985.

[5] S. Blom. An approximation based approach to infinitary lambda calculi. In *Proceedings of the 15th International Conference on Rewriting Techniques and Applications (RTA 2004)*, volume 3091 of *Lecture Notes in Computer Science*, pages 221–232. Springer-Verlag, 2004.

[6] B. Courcelle. Fundamental properties of infinite trees. *Theoretical Computer Science*, 25(2):95–169, 1983.

[7] N. Dershowitz and S. Kaplan. Rewrite, rewrite, rewrite, rewrite, rewrite. In *Proceedings of the 16th Annual ACM Symposium on Principles of Programming Languages (POPL '89)*, pages 250–259, 1989.

[8] N. Dershowitz, S. Kaplan, and D. A. Plaisted. Infinite normal forms. In *Proceedings of the 16th International Colloquium on Automata, Languages and Programming (ICALP '89)*, volume 372 of *Lecture Notes in Computer Science*, pages 249–262. Springer-Verlag, 1989.

[9] N. Dershowitz, S. Kaplan, and D. A. Plaisted. Rewrite, rewrite, rewrite, rewrite, rewrite, ... *Theoretical Computer Science*, 83(1):71–96, 1991.

[10] M. Dezani-Ciancaglini, P. Severi, and F.-J. de Vries. Infinitary lambda calculus and discrimination of Berarducci trees. *Theoretical Computer Science*, 298(2):275–302, 2003.

[11] W. M. Farmer and R. J. Watro. Redex capturing in term graph rewriting. In *Proceedings of the 4th Conference on Rewriting Techniques and Applications (RTA '91)*, volume 488 of *Lecture Notes in Computer Science*, pages 13–24, 1991.

[12] J. Glauert, D. Kesner, and Z. Khasidashvili. Expression Reduction Systems and extensions: An overview. In *Processes, Terms and Cycles: Steps on the Road to Infinity*, volume 3838 of *Lecture Notes in Computer Science*, pages 496–553. Springer-Verlag, 2005.

[13] J. A. Goguen, J. W. Thatcher, E. G. Wagner, and J. B. Wright. Initial algebra semantics and continuous algebras. *Journal of the ACM*, 24(1):68–95, 1977.

[14] M. Hanus and C. Prehofer. Higher-order narrowing with definitional trees. In *Proceedings of the 7th International Conference on Rewriting Techniques and Applications (RTA '96)*, volume 1103 of *Lecture Notes in Computer Science*, pages 138–152. Springer-Verlag, 1996.

[15] S. Kahrs. Compilation of combinatory reduction systems. In *Proceedings of the 1st International Workshop on Higher-Order Algebra, Logic, and Term Rewriting (HOA '93)*, volume 816 of *Lecture Notes in Computer Science*, pages 169–188. Springer-Verlag, 1993.

[16] J. R. Kennaway, J. W. Klop, M. R. Sleep, and F.-J. de Vries. Infinitary lambda calculus. *Theoretical Computer Science*, 175(1):93–125, 1997.

[17] R. Kennaway and F.-J. de Vries. Infinitary rewriting. In Terese [35], Chapter 12.

[18] R. Kennaway, J. W. Klop, R. Sleep, and F.-J. de Vries. Transfinite reductions in orthogonal term rewriting systems. *Information and Computation*, 119(1):18–38, 1995.

[19] R. Kennaway, V. van Oostrom, and F.-J. de Vries. Meaningless terms in rewriting. *The Journal of Functional and Logic Programming*, 1, 1999.

[20] J. Ketema and J. G. Simonsen. Infinitary combinatory reduction systems. In *Proceedings of the 16th International Conference on Rewriting Techniques and Applications (RTA 2005)*, volume 3467 of *Lecture Notes in Computer Science*, pages 438–452. Springer-Verlag, 2005.

[21] J. Ketema and J. G. Simonsen. On confluence of infinitary combinatory reduction systems. In *Proceedings of the 12th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR 2005)*, volume 3835 of *Lecture Notes in Artificial Intelligence*, pages 199–214. Springer-Verlag, 2005.

[22] J. W. Klop. *Combinatory Reduction Systems*. PhD thesis, Rijksuniversiteit Utrecht, 1980.

[23] J. W. Klop and R. de Vrijer. Infinitary normalization. In S. N. Artëmov, H. Barringer, A. S. d'Avila Garcez, L. C. Lamb, and J. Woods, editors, *We Will Show Them: Essays in Honour of Dov Gabbay*, volume 2, pages 169–192. College Publications, 2005.

[24] J. W. Klop, V. van Oostrom, and F. van Raamsdonk. Combinatory reduction systems: introduction and survey. *Theoretical Computer Science*, 121(1 & 2):279–308, 1993.

[25] B. Lisper. Infinite unfolding and transformations of nondeterministic programs. *Fundamenta Informaticae*, 66(4):415–439, 2005.

[26] M. Makkai. Admissible sets and infinitary logic. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 233–282. North-Holland, 1977.

[27] A. Middeldorp. Call by need computations to root-stable form. In *Proceedings of the 24th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL '97)*, pages 94–105, 1997.

[28] T. Nipkow, L. C. Paulson, and M. Wenzel. *Isabelle/HOL: A Proof Assistant for Higher-Order Logic*, volume 2283 of *Lecture Notes in Computer Science*. Springer-Verlag, 2002.

[29] T. Nipkow and C. Prehofer. Higher-order rewriting and equational reasoning. In W. Bibel and P. Schmitt, editors, *Automated Deduction – A Basis for Applications*, volume 1, pages 399–430. Kluwer, 1998.

[30] M. J. O'Donnell. *Computing in Systems Described by Equations*, volume 58 of *Lecture Notes in Computer Science*. Springer-Verlag, 1977.

[31] M. J. Plasmeijer and M. C. J. D. van Eekelen. *Functional Programming and Parallel Graph Rewriting*. Addison-Wesley, 1993.

[32] R. C. Sekar and I. V. Ramakrishnan. Programming in equational logic: beyond strong sequentiality. *Information and Computation*, 104(1):78–109, 1993.

[33] P. Severi and F.-J. de Vries. An extensional Böhm model. In *Proceedings of the 13th International Conference on Rewriting Techniques and Applications (RTA 2002)*, volume 2378 of *Lecture Notes in Computer Science*, pages 159–173. Springer-Verlag, 2002.

[34] J. G. Simonsen. On modularity in infinitary rewriting. *Information and Computation*, 204(6):957–988, 2006.

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

[36] The Coq Development Team. The Coq proof assistant reference manual, version 8.0. Technical report, INRIA, 2006.

[37] V. van Oostrom. Higher-order families. In *Proceedings of the 7th International Conference on Rewriting Techniques and Applications (RTA '96)*, volume 1103 of *Lecture Notes in Computer Science*, pages 392–407. Springer-Verlag, 1996.

[38] V. van Oostrom. Normalisation in weakly orthogonal rewriting. In *Proceedings of the 10th International Conference on Rewriting Techniques and Applications (RTA '99)*, volume 1631 of *Lecture Notes in Computer Science*, pages 60–74. Springer-Verlag, 1999.

[39] F. van Raamsdonk. Outermost-fair rewriting. In *Proceedings of the 3rd International Conference on Typed Lambda Calculi and Applications (TLCA '97)*, volume 1210 of *Lecture Notes in Computer Science*, pages 284–299. Springer-Verlag, 1997.

[40] F. van Raamsdonk. Higher-order rewriting. In Terese [35], Chapter 11.

A Proof of Proposition 6.6 and Lemma 6.7

As in Section 6, we assume in this appendix that we are working in an orthogonal iCRS and that $\mathcal{U}$ is a set of redexes in a term s. We first prove Proposition 6.6.

Proof of Proposition 6.6. As each path projection derives from a path, we have by definition that ϕ is *surjective*. Similar for the path projections in $\mathcal{P}(s,\mathcal{U})$ and the maximal paths, as each path projection in $\mathcal{P}(s,\mathcal{U})$ derives from a maximal path.

To prove that ϕ is *injective*, suppose there exist (maximal) paths Π, Π' such that $\phi(\Pi) = \phi(\Pi')$. By definition of ϕ both paths and the path projection consist of the same number of nodes and edges. Let Π^* be the longest shared prefix of Π and Π'. The prefix Π^* is non-empty, as any path of s starts with (s, ϵ). There are now two cases to consider depending on Π^* ending in either an edge or a node.

In case Π^* ends in an edge, the next node is uniquely determined by the definition of paths. Hence, as Π and Π' have the same number of nodes and edges we can extend Π^* with that unique node, contradiction.

In case Π^* ends in a node, both paths extend Π^*, otherwise $\Pi = \Pi'$ or the paths differ in the number of nodes or edges. In case the extension is with an unlabelled edge in case of one of the paths, the other path must also extend Π^* with an unlabelled edge. This follows by the definition of paths. In case the extension is with an edge labelled i, the other path must also extend Π^* with an edge labelled i. This follows by definition of paths and by $\phi(\Pi) = \phi(\Pi')$. Hence, in case Π^* ends in a node a contradiction also follows. We can conclude that ϕ is an injection both between paths and path projections and between maximal paths and the path projections in $\mathcal{P}(s,\mathcal{U})$. □

To prove Lemma 6.7, we define a map θ_u taking maximal paths Π of s with respect to $\mathcal{U}$ to maximal paths of t with respect to $\mathcal{U}/u$, where $u \in \mathcal{U}$ and $s \to t$ by contracting u. The definition of $\theta_u(\Pi)$ employs a partial map ψ that has three arguments: a node of Π, a finite string over $\mathbb{N}$, and a partial map from $\mathcal{U} - \{u\}$ to finite strings over $\mathbb{N}$.

We first define ψ. In the definition, given a partial map ρ, we denote by $\rho[x \mapsto y]$ the partial map ρ' defined as:

$$\rho'(z) = \begin{cases} y & \text{if } z = x \\ \rho(z) & \text{otherwise} \end{cases}$$

Definition A.1. Let Π be a maximal path of s with respect to $\mathcal{U}$, let $u \in \mathcal{U}$, and let $s \to t$ by contracting u. Define $\psi(n, q_t, \rho)$ as:

1. If n is labelled (s, p) with the subterm at p not a redex in $\mathcal{U}$, then
   (a) if n has no outgoing edge define $\psi(n, q_t, \rho) = (t, q_t)$,
   (b) if n has an edge labelled i to n' define $\psi(n, q_t, \rho) = (t, q_t) \xrightarrow{i} \psi(n', q_t \cdot i, \rho)$.

2. If n is labelled (s, p_v) with $v \in \mathcal{U} - \{u\}$ and if n has an unlabelled edge to n', then define $\psi(n, q_t, \rho) = (t, q_t) \to \psi(n', q_t, \rho[v \mapsto q_t])$,
3. If n is labelled (s, p_u) and if n has an unlabelled edge to n', then define $\psi(n, q_t, \rho) = \psi(n', q_t, \rho)$.
4. If n is labelled (s, p) with $s|_p$ a variable bound by $v \in \mathcal{U} - \{u\}$ and if n has an unlabelled edge to n', then define $\psi(n, q_t, \rho) = (t, q_t) \to \psi(n', \rho(v), \rho)$.
5. If n is labelled (s, p) with $s|_p$ a variable bound by u and if n has an unlabelled edge to n', then define $\psi(n, q_t, \rho) = \psi(n', q_t, \rho)$.
6. If n is labelled (r, p, p_v) with $r|_p$ not a meta-variable and $v \in \mathcal{U} - \{u\}$, then
   (a) if n has no outgoing edge define $\psi(n, q_t, \rho) = (r, p, q_t)$,
   (b) if n has an edge labelled i to n' define $\psi(n, q_t, \rho) = (r, p, q_t) \xrightarrow{i} \psi(n', q_t, \rho)$.
7. If n is labelled (r, p, p_u) with $r|_p$ not a meta-variable, then
   (a) if n has no outgoing edge define $\psi(n, q_t, \rho) = (t, q_t)$,
   (b) if n has an edge labelled i to n' define $\psi(n, q_t, \rho) = (t, q_t) \xrightarrow{i} \psi(n', q_t \cdot i, \rho)$.
8. If n is labelled (r, p, p_v) with $r|_p$ a meta-variable and $v \in \mathcal{U} - \{u\}$ and if n has an unlabelled edge to n', which is labelled $(s, p_v \cdot q)$, then define $\psi(n, q_t, \rho) = (r, p, q_t) \to \psi(n', q_t \cdot q, \rho)$.
9. If n is labelled (r, p, p_u) with $r|_p$ a meta-variable and if n has an unlabelled edge to n', which is labelled $(s, p_u \cdot q)$, then define $\psi(n, q_t, \rho) = \psi(n', q_t, \rho)$.

Let $\perp$ be completely undefined map. We define the following.

Definition A.2. Let $u \in \mathcal{U}$ and let Π be a maximal path of s with respect to $\mathcal{U}$. The map θ_u is defined as:

$$\theta_u(\Pi) = \psi((s, \epsilon), \epsilon, \perp)\,.$$

Note that $\theta_u(\Pi)$ is calculated by iteration of ψ. After a finite number of iterations, a finite prefix of $\theta_u(\Pi)$ is obtained.

We next show that that θ_u is well-defined: $\theta_u(\Pi)$ is a maximal path of t with respect to $\mathcal{U}/u$.

Proposition A.3. *Let Π be a maximal path of s with respect to $\mathcal{U}$ and let $u \in \mathcal{U}$. For each finite number of iterations of ψ in the calculation of $\theta_u(\Pi)$ the following holds:*

- *Either no nodes and edges have been generated, or the generated nodes and edges, with exception of the edge generated last, form a path, and the edge generated last is a valid one when extending the path to a longer one.*

- *For all defined $\rho(v)$ in the third argument of ψ it holds that $\rho(v) \in \mathcal{P}os(t)$ and that a residual of v occurs at $\rho(v)$ in t.*

Distinguishing on the particular clause of Definition A.1 employed in the last iteration, the following also holds:

(1) *(a) nothing; (b) q_t is descendant of p and the next node generated is $(t, q_t \cdot i)$, which together with the previously generated nodes and edges forms a path;*

(2) *a residual of v occurs at q_t in t and the next node generated is (r, ϵ, q_t), which together with the previously generated nodes and edges forms a path;*

(3) *$q_t = p_u$ and the next node generated is (t, q_t), which together with the previously generated nodes and edges forms a path;*

(4) *if n' is labelled $(r, p' \cdot i, p_v)$, then the next node generated is $(r, p' \cdot i, \rho(v))$, which together with the previously generated nodes and edges forms a path;*

(5) *there are two subcases:*

   - *if the previous iterations employ clause (3) followed by a number of iterations employing in turn clauses (5) and (9), then $q_t = p_u$;*
   - *if the previous iteration employs clause (1) or if the previous iterations employ clause (7) followed by a number of iterations employing in turn clauses (5) and (9), then $q_t = q'_t \cdot i$ where q'_t is the q_t from either clause (1) or (7);*

   in both cases the next node generated is (t, q_t), which together with the previously generated nodes and edges forms a path.

(6) *(a) nothing; (b) a residual of v occurs at q_t and the next node generated is $(r, p \cdot i, q_t)$, which together with the previously generated nodes and edges forms a path;*

(7) *(a) nothing; (b) $q_t = p_u \cdot q$ where q is a descendant of p across a complete development of the parallel β-redexes in r_σ, with r_σ as defined above Definition 4.14, and the next node generated is $(t, q_t \cdot i)$, which together with the previously generated nodes and edges forms a path;*

(8) *q_t is a descendant of p_v and the next node generated is $(t, q_t \cdot q)$, which together with the previously generated nodes and edges forms a path;*

(9) *there are two subcases:*

   - *if the next iteration does* not *employ clause (5), then q_t is a descendant of $p_u \cdot q$;*
   - *otherwise, q_t is as in clause (5);*

   in both cases the next node generated is (t, q_t), which together with the previously generated nodes and edges forms a path.

In addition, if Π *is a maximal path of* s *with respect to* $\mathcal{U}$*, then* $\theta_u(t)$ *is a maximal path of* t *with respect to* $\mathcal{U}/u$.

Proof. Let Π be a maximal path of s with respect to $\mathcal{U}$, let $u \in \mathcal{U}$, and let $s \to t$ by contracting u. We prove the lemma by induction on the number of iterations of ψ.

Below, when we say that something is a path of t, we implicitly assume that it is a path of t with respect to $\mathcal{U}/u$.

Base case. By definition of ψ we have that only clauses (1), (2), and (3) can be employed in the first iteration. The other clauses either require a bound variable at the root of s, which is impossible, or they require the label of the node in the first argument to be a triple, which is not the case. We deal with each of the possible clauses in turn:

(1) In this case a node labelled (t, ϵ) is generated and possibly an edge labelled i. Obviously, (t, ϵ) is a path. The partial map $\perp$ is unaffected by this clause, thus satisfying the necessary requirements.

As the current iteration implies that no redex from $\mathcal{U}$ occurs at the root of s, we have that $q_t = \epsilon \in \mathcal{P}os(t)$ is a descendant of $p = \epsilon$ and $root(t|_\epsilon) = root(s|_\epsilon)$. Hence, in case of clause (a), the path is maximal like Π. In case of clause (b), the edge labelled i is allowed and the next node generated node must be (t, i). This node forms a path together with (t, ϵ) and the edge labelled i. By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(2) In this case a node labelled (t, ϵ) and an unlabelled edge are generated. Obviously, (t, ϵ) is a path. As orthogonality is assumed, a redex v', which is a residual of v, occurs at $\epsilon \in \mathcal{P}os(t)$. Hence, $\perp[v \mapsto \epsilon]$ satisfies the necessary requirements.

As $v \in \mathcal{U} - \{u\}$, it holds that $v' \in \mathcal{U}/u$. Hence, the unlabelled edge is allowed, and the next node generated must be (r, ϵ, ϵ), where r is the right-hand side of the rewrite rule employed in v'. This node forms a path together with (t, ϵ) and the unlabelled edge. By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(3) Obviously, $\perp$ is unaffected by this clause, thus satisfying the necessary requirements. Moreover, $q_t = \epsilon \in \mathcal{P}os(t)$ is equal to $p = \epsilon$, and by definition of ψ the next generated node must be (t, ϵ), which is a path. By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

Induction step. Assume we have proved the lemma up to some arbitrary number of iterations. We next prove that it also holds in case of one more iteration. We deal with each of the possible clauses in turn:

(1) In this case the only possible clauses employed in the previous iteration are (1), (8), and (9). All other clauses force the label of the node in the first argument of ψ to be a triple, which is not the case.

A node labelled (t, q_t) is generated and possibly an edge labelled i. By the clauses possible in the previous iteration, (t, q_t) forms a path together with the previously generated nodes and edges. The partial map ρ is unaffected by this clause, thus satisfying the necessary requirements.

Also by the clauses possible in the previous iteration, $q_t \in \mathcal{P}os(t)$ is a descendant of p and $root(t|_{q_t}) = root(s|_p)$. Hence, in case of clause (a), the path is maximal like Π. In case of clause (b), the edge labelled i is allowed and the next node generated must be $(t, q_t \cdot i)$. This node forms a path together with previously generated nodes and edges. By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(2) As before, the only possible clauses employed in the previous iteration are (1), (8), and (9). All other clauses force the label of the node in the first argument of ψ to be a triple, which is not the case.

A node labelled (t, q_t) and an unlabelled edge are generated. By the clauses possible in the previous iteration, (t, q_t) forms a path together with the previously generated nodes and edges. Moreover, as orthogonality is assumed, a redex v', which is a residual of v, occurs at $q_t \in \mathcal{P}os(t)$. Hence, as ρ satisfies the necessary requirements, $\rho[v \mapsto q_t]$ does so too.

As $v \in \mathcal{U} - \{u\}$, it holds that $v' \in \mathcal{U}/u$. Hence, the unlabelled edge is allowed, and the next node generated is (r, ϵ, q_t), where r is the right-hand side of the rewrite rule employed in v'. This node forms a path together with previously generated nodes and edges. By the construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(3) In this case the only possible clauses employed in the previous iteration are (1) and (8). All other clauses, except (9), force the label of the node in the first argument of ψ to be a triple, which is not the case. Clause (9) is impossible as it requires the redex u to occur above itself in s.

Obviously, ρ is unaffected by this clause, thus satisfying the necessary requirements. By the clauses possible in the previous iteration and the definition of descendants, $q_t = p_u$. Also by the clauses possible in the previous iteration, the next node generated is (t, q_t) and the node forms a path together with previously generated nodes and edges. By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(4) As before, the only possible clauses employed in the previous iteration are (1), (8), and (9). All other clauses force the label of the node in the first argument of ψ to be a triple, which is not the case.

A node labelled (t, q_t) and an unlabelled edge are generated. By the clauses possible in the previous iteration, (t, q_t) forms a path together with the previously generated nodes and edges. The partial map ρ is unaffected by this clause, thus satisfying the necessary requirements.

Also by the the clauses possible in the previous iteration, q_t is a descendant of p and $root(t|_{q_t})$ is a variable bound by a residual v' of v in t, where by construction $\rho(v)$ is the position of v'. Hence, the unlabelled edge is allowed. That a node labelled $(r, p', \rho(v))$, where r is the right-hand side of the rewrite rule employed in v', has been generated as the last node labelled with $\rho(v)$ follows by definition of ψ. Hence, the next node generated is $(r, p \cdot i, \rho(v))$. This node forms a path together with previously generated nodes and edges. By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(5) As before, the only possible clauses employed in the previous iteration are (1), (8), and (9). All other clauses force the label of the node in the first argument of ψ to be a triple, which is not the case.

Obviously, ρ is unaffected by this clause, thus satisfying the necessary requirements. By the clauses in the previous iteration, the requirements of the two subcases are satisfied and the next node generated is (t, q_t). By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(6) In this case the only possible clauses employed in the previous iteration are (2), (4), and (6). Clauses (1), (8), and (9) force the label of the node in the first argument of ψ to be a tuple, which is not the case. Clauses (3) and (5) force v to be equal to u, which is not allowed.

A node labelled (r, p, q_t) is generated and possibly an edge labelled i. By the clauses possible in the previous iteration, (r, p, q_t) forms a path together with the previously generated nodes and edges. The partial map ρ is unaffected by this clause, thus satisfying the necessary requirements.

As r is left unchanged, it holds in case of clause (a), that the path is maximal like Π. In case of clause (b), the edge labelled i is allowed and the next node generated is $(r, p \cdot i, q_t)$. This node forms a path together with previously generated nodes and edges. By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(7) In this case the only possible clauses employed in the previous iteration are (3), (5), and (7). Clauses (1), (8), and (9) force the label of the node in the first argument of ψ to be a tuple, which is not the case. Clauses (2) and (4) force v to be unequal to u.

A node labelled (t, q_t) is generated and possibly an edge labelled i. By the clauses possible in the previous iteration (t, q_t) forms a path together with the previously generated nodes and edges. The partial map ρ is unaffected by this clause, thus satisfying the necessary requirements.

Also by the clauses possible in the previous iteration, $q_t = p_u \cdot q$ and $root(t|_{q_t}) = root(r_\sigma|_p)$, where q is a descendant of p across complete development of the parallel β-redexes in r_σ, with r_σ as defined above Definition 4.14. Hence, in case of clause (a), the path is maximal like Π. In case of clause (b), the edge labelled i is allowed. Moreover, the next node generated is $(t, q_t \cdot i)$. This node forms a path together with previously generated nodes and edges. By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(8) In this case the only possible clauses employed in the previous iteration are (2), (4), and (6). Clauses (1), (8), and (9) force the label of the node in the first argument of ψ to be a tuple, which is not the case. Clauses (3) and (5) force v to be equal to u, which is not allowed.

In this case a node labelled (r, p, q_t) and an unlabelled edge are generated. By the clauses possible in the previous iteration, (r, p, q_t) forms a path together with the previously generated nodes and edges. The partial map ρ is unaffected by this clause, thus satisfying the necessary requirements.

As r is left unchanged, the unlabelled edge is allowed. Moreover, as a residual of v occurs at q_t, the next node generated is $(t, q_t \cdot q)$. This node forms a path together with previously generated nodes and edges. By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

(9) In this case the only possible clauses employed in the previous iteration are (3), (5), and (7). Clauses (1), (8), and (9) force the label of the node in the first argument of ψ to be a tuple, which is not the case. Clauses (2) and (4) force v to be unequal to u.

Obviously, ρ is unaffected by this clause, thus satisfying the necessary requirements. By the clauses in the previous iteration, the requirement that q_t is a descendant of $p_u \cdot q$ is satisfied in the first subcase, and the requirements of clause (5) are satisfied in the second subcase. Moreover, the next node generated is (t, q_t). By construction of paths, the finite prefix of Π considered thus far is not maximal and there is nothing more to show.

There can be no infinite cycle of iterations employing clauses (5) and (9) in the construction of $\theta_u(\Pi)$, because that implies the existence of an infinite chain of meta-variables in r. Hence, $\theta_u(\Pi)$ is a well-defined path of t with respect to $\mathcal{U}/u$. In case Π is finite, the induction shows that $\theta_u(\Pi)$ is maximal. In case Π is infinite, so is $\theta_u(\Pi)$, and we conclude that $\theta_u(\Pi)$ is a maximal path. □

The next lemma relates the maximal paths of s with respect to $\mathcal{U}$ to the maximal paths of t with respect to $\mathcal{U}/u$. In the proof of the lemma we leave out the labels of the explicitly denoted edges.

Proposition A.4. *The map θ_u is a bijection.*

Proof. Let $u \in \mathcal{U}$ and $s \to t$ by contracting u. By Proposition A.3, we have that θ_u maps maximal paths of s with respect to $\mathcal{U}$ to maximal paths of t with respect to $\mathcal{U}/u$.

To prove that θ_u is *surjective*, let Π_t be a maximal path of t with respect to $\mathcal{U}/u$. We are done if $\Pi_t = \theta_u(\Pi_s)$ for some maximal path Π_s of s with respect to $\mathcal{U}$. Otherwise, Π_t has a finite non-empty prefix Π'_t in common with $\theta_u(\Pi_s)$ for some maximal path Π_s of s with respect to $\mathcal{U}$. The prefix is non-empty since any path of t begins with (t, ϵ). Let Π'_t be the longest finite prefix of Π_t such that $\theta_u(\Pi_s) = \Pi'_t \to \cdots$ for some maximal path Π_s. We have $\Pi_s = \Pi'_s \to \cdots$ for some finite path Π'_s. By definition of θ_u we can extend the prefix Π'_t with a new node precisely when we can extend Π'_s. Hence, Π'_t cannot be the *longest* finite prefix with $\theta_u(\Pi_s) = \Pi'_t \to \cdots$ for some maximal path Π_s in s, since we can extend Π'_s to form a new maximal path with more nodes, contradiction. Hence, $\Pi_t = \theta_u(\Pi_s)$ for some maximal path Π_s.

To prove that θ_u is *injective*, suppose there exist two maximal paths Π and Π' of s with respect to $\mathcal{U}$ such that $\theta_u(\Pi) = \theta_u(\Pi')$. Let Π^* be the longest prefix shared between Π and Π'. The prefix Π^* is non-empty, as any path of s begins with (s, ϵ). There are now two cases to consider depending on Π^* ending in an edge or a node.

In case Π^* ends in an edge, the next node is uniquely determined by the definition of paths. Hence, as Π and Π' are maximal, we can extend Π^* with that unique node, contradiction.

In case Π^* ends in a node, at least one of Π and Π' extends Π^*, otherwise $\Pi = \Pi'$. In case the extension is with an unlabelled edge, the other path must also extend Π^* with an unlabelled edge. This follows by the definition of paths and by Π and Π' being maximal. Otherwise, in case the extension is with an edge labelled i, the other path must also extend Π^* with an edge labelled i. This follows by definition of paths and $\theta_u(\Pi) = \theta_u(\Pi')$. Hence, in case Π^* ends in a node a contradiction also follows and we can conclude that θ_u is injective. □

We finally prove Lemma 6.7.

Proof of Lemma 6.7. By Proposition A.4, the map θ_u is a bijection between the maximal paths of s with respect to $\mathcal{U}$ and the maximal paths of t with respect to $\mathcal{U}/u$. By Proposition 6.6, a bijection exists between the set of paths and the set of path projections mapping unlabelled edges to ϵ-labelled edges and labelled edges to edges with the same label. Hence, θ_u induces a bijection θ'_u between $\mathcal{P}(s,\mathcal{U})$ and $\mathcal{P}(t,\mathcal{U}/u)$. By examining the construction of θ_u, we see that it only deletes unlabelled edges and nodes corresponding to meta-variables of u and variables bound by u. Moreover, it is evident that if an infinite sequence of nodes and unlabelled edges were deleted, the right-hand side of the rule of u would contain an infinite chain of meta-variables, contradicting the definition of meta-terms. Hence, $\phi(\theta'_u(\Pi))$ can be obtained by $\phi(\Pi)$ by deleting only finite sequences of unlabelled nodes and ϵ-labelled edges, as required. □