Consiglio Nazionale delle Ricerche

ISTITUTO DI ELABORAZIONE DELLA INFORMAZIONE

PISA

Infinite Normal Forms for Non Linear Term Rewriting Systems

P. Inverardi, M. Nesi

Nota Interna IEI B4-41

Settembre 1990

Infinite Normal Forms for Non Linear Term Rewriting Systems

P. Inverardi, M. Nesi
IEI - CNR, via S. Maria 46, I-56126 Pisa, ITALY

Extended Abstract

1. Introduction

In this paper we discuss the applicability of the framework defined in [DK89, DKP89, DKP90] on the existence of infinite normal forms on a particular class of non left linear term rewriting system.

It is very common, in practice, to deal with non linear term rewriting systems especially when one wants to axiomatize the semantics of a given structure, and the purpose of this paper is that of showing that under some hypotheses it is possible to extend the results of [DK89, DKP89, DKP90] to a class of non left linear term rewriting system. Although, at a first glance the class we consider can appear quite restrictive, in practice it captures a number of algebraic structures largely used in computer science.

In particular, we deal with term rewriting systems in which the non terminating rules are *unfolding* rules that model the operational semantics of a recursive operator. The left linearity requirement is replaced by the *retraction* property of the supporting term algebra that allows the definition of a rewriting relation modulo an equivalence relation induced on the set of terms by the unfolding rules. With these two assumptions we can still restrict to consider, as in [DK89, DKP89, DKP90], only a subset of infinite derivations, i.e. fair derivations. Actually, we go further on by focussing on those rewriting systems which admit a peculiar kind of fair derivations, i.e. *uniform* systems and *structured* fair derivations. The ω-confluence of the rewriting system can then be proved by properly constraining the possible interaction between the non terminating rules and the remaining rules. In this respect the notion of *independence* on the rules of the rewriting system is introduced.

The approach has been used in [IN90b] to prove the existence of infinite normal forms for recursive (finite state) CCS expressions [Mil80] with respect to a correct and complete axiomatization of the observational congruence given by Milner [Mil89]. In fact, our interest in non terminating non linear rewriting systems comes from the experience we have made by developing a verification system for the CCS language based on term rewriting techniques [DIN90, IN90a]. In that framework it results that all the axiomatic characterizations of the various behavioural equivalences contain non left linear rules. On the other hand, non termination arises as soon as one wants to consider recursive processes.

2. Basic Definitions

We assume that the reader is familiar with the basic concepts of term rewriting systems. We summarize the most relevant definitions below, while we refer to [BD89, DK89, DKP89] for more details.

Let Σ be a set of operators, V be a set of variables and $T_\Sigma(V)$ denotes the set of terms over Σ and V. An *equational theory* is any set $E = \{(s, t) \mid s, t \in T_\Sigma(V)\}$. Elements (s, t) are called *equations* and written s = t. Let $\sim_E$ be the smallest symmetric relation that contains E and is closed under monotonicity and substitution. Let $=_E$ be the reflexive-transitive closure of $\sim_E$. Given Σ, V and an equational theory S, $T_\Sigma(V)_{/S}$ denotes the *quotient algebra* with respect to S.

A *term rewriting system* (TRS) R is any set $\{(l_i, r_i) \mid l_i, r_i \in T_\Sigma(V), \mathcal{V}(r_i) \subseteq \mathcal{V}(l_i)\}$. The pairs (l_i, r_i) are called *rewriting rules* and written $l_i \to r_i$. The *rewriting relation* $\to_R$ on $T_\Sigma(V)$ is defined as the smallest relation containing R that is closed under monotonicity and substitution. A term t *rewrites* to a term s, written $t \to_R s$, if there exist $l_k \to r_k$ in R, a substitution σ and a subterm t/u at the occurrence u, called *redex*, such that $t/u = l_k\sigma$ and $s = t[u \leftarrow r_k\sigma]$. A TRS R is *left linear* if the left hand side l of each rule $l \to r$ in R has at most one occurrence of any variable. A term t is said to *overlap* a term t' if t unifies with a non variable subterm of t' (after renaming the variables in t so as not to conflict with those in t').

Let $\xrightarrow{+}$ and $\xrightarrow{*}$ denote the transitive and transitive-reflexive closure of $\to$, respectively.

A TRS R is *finitely terminating* if there is no infinite sequence $t_1 \to_R t_2 \to_R \ldots$. A TRS R is *confluent* if whenever $s \;{}_R\!\xleftarrow{*}\; t \xrightarrow{*}_R q$, then there exists a term t' such that $s \xrightarrow{*}_R t' \;{}_R\!\xleftarrow{*}\; q$, while R is *locally confluent* if whenever $s \;{}_R\!\leftarrow t \to_R q$, then there exists a term t' such that $s \xrightarrow{*}_R t' \;{}_R\!\xleftarrow{*}\; q$. A term t is *in R-normal form* if there exists no term s such that $t \to_R s$. A term s is a *R-normal form* of t if $t \xrightarrow{*}_R s$ and s is in R-normal form. A TRS R is *canonical* if it is finitely terminating and confluent.

An *equational* TRS is a tuple (R, E), written R/E, where R is a TRS and E an equational theory. The rewriting relation $\to_{R/E}$ is defined by $=_E \bullet \to_R \bullet =_E$, where $\bullet$ denotes composition of relations.

Given a (possibly infinite) rewriting relation $\to$, let us recall the following definitions:

Definition 1 (ω-rewriting) $t \to^\omega t'$ if $t \xrightarrow{*} t'$ or if there exists an infinite *derivation* $t = t_0 \to t_1 \to \ldots \to t_n \to \ldots$ such that $\lim_{n\to\infty} t_n = t'$.

Definition 2 (ω-terminating) $\to$ is ω-*terminating* if for any infinite derivation $t = t_0 \to t_1 \to \ldots \to t_n \to \ldots$ of terms, the limit $\lim_{n\to\infty} t_n$ exists.

Definition 3 (top-terminating) $\rightarrow$ is *top-terminating* if there are no infinite derivations $t = t_0 \rightarrow t_1 \rightarrow \ldots \rightarrow t_n \rightarrow \ldots$ of terms with infinitely many rewrites at the topmost occurrence.
Definition 4 (ω-confluence) $\rightarrow$ is ω-*confluent* if ${}^{\omega}\!\!\leftarrow \bullet \rightarrow^{\omega}$ implies $\rightarrow^{\omega} \bullet {}^{\omega}\!\!\leftarrow$.
In other words, for any t, t_1, t_2 such that $t_1 \;{}^{\omega}\!\!\leftarrow t \rightarrow^{\omega} t_2$, there exists t' such that $t_1 \rightarrow^{\omega} t' \;{}^{\omega}\!\!\leftarrow t_2$.
Definition 5 (ω-canonicity) $\rightarrow$ is ω-*canonical* if it is ω-terminating and ω-confluent.
Definition 6 (ω-normal form) A term t' is an ω-*normal form* of t if $t \rightarrow^{\omega} t'$ and t' is minimal for $\rightarrow$, i.e. if $t' \rightarrow t''$, then $t'' = t'$.
Thus, an ω-normal form need not be irreducible.
Among all the possible infinite derivations from any term, we can single out "interesting" derivations.
Definition 7 (fair derivation [DKP90]) A derivation $t_0 \rightarrow t_1 \rightarrow \ldots \rightarrow t_n \rightarrow \ldots$ is *fair* if whenever there is a rule $l \rightarrow r$ and an occurrence u such that, for all n past some N, the subterm t_n/u is a redex for $l \rightarrow r$, then (at least) one of the rule applications $t_n \rightarrow t_{n+1}$ ($n \geq N$) is an application of $l \rightarrow r$ at u.

Thus, a fair derivation guarantees that a redex does not persist forever. Note that this definition does not prevent the fact that the same rewriting rule is applicable infinitely many times at deeper and deeper occurrences.

3. Non Left Linear Term Rewriting System and ω-Normal Forms

Now, the interesting point is to show that fair derivations are the only derivations we have to look at. This means that the limit of a fair derivation can be shown to be an ω-normal form and, viceversa, the ω-normal form of any term can be computed as the limit of a fair derivation. This result has been shown in [DK89, DKP89, DKP90] by using the further hypothesis of left linearity of the term rewriting system. On the other hand, we will discuss what happens when the left linearity hypothesis is not satisfied and which new hypotheses have to be assumed on an infinite rewriting relation in order to guarantee the existence of ω-normal forms.

Definition 8 (unfolding) An *unfolding* is an equation $G = H[G]$, where G is a non variable subterm of t_i, $1 \leq i \leq k$, in the context $H[\ ] = op(t_1, \ldots, t_k)$ for some $op \in \Sigma$ of arity k.
Definition 9 (unfolding rule) Given an unfolding $G = H[G]$, the non terminating rule $G \rightarrow H[G]$ is an *unfolding rule*.

Definition 10 (retraction) Let S be a set of unfoldings. $T_\Sigma(V)_{/S}$ is *retractile by* $=_F$ if it is possible to define on it an equivalence relation $=_F$ such that for any two terms $t_1, t_2 \in T_\Sigma(V)_{/S}$, $t_1 =_F t_2$ implies the following:

i. it exists a term t' such that $t_1 \rightarrow_S{}^\omega\ t'\ {}^\omega{}_S\!\leftarrow t_2$;

ii. for each equivalence class of F there exists a unique finite canonical representative.

The retraction property means that our signature is such that any infinite term t can be seen as the infinite unfolding of a finite term. Furthermore, if a class of finite terms exists whose unfolding results in the same infinite term t, it is possible to select a unique finite canonical representative of the class, C(t). Thus, an equivalence relation on $T_\Sigma(V)_{/S}$, denoted with $=_F$, can be determined such that for any two infinite terms t_1, t_2, the equivalence $t_1 =_F t_2$ holds if and only if $C(t_1) = C(t_2)$.

Thus, let us now assume that we are dealing with term rewriting systems T, defined on a retractile $T_\Sigma(V)_{/S}$ equipped with $=_F$, such that T is the union $R/F \cup S$ of a finitely terminating term rewriting system R and a term rewriting system S which only contains unfolding rules.

Example 1

R	$x + e \rightarrow x$
	$x + x \rightarrow x$
S	$f(x) \rightarrow g(f(x))$

The term rewriting system $T = R/F \cup S$ is not finitely terminating and not left linear. R is a finitely terminating term rewriting system and S contains only one non terminating rule. The equivalence $=_F$ is trivially defined by the unfolding $f(x) = g(f(x))$, which collapses all the terms like $g^n(f(t))$ to the canonical representative f(x) for any n and t, and consequently the canonical representative of all the other terms can be obtained.

In all the following definitions and propositions it is assumed to deal with the above characterized term rewriting systems T. Now we are going to discuss the ω-canonicity of such systems; this means that we discuss when $\rightarrow_T$ is:

- top-terminating;
- ω-terminating;
- ω-confluent.

We first show that $\rightarrow_T$ is top-terminating and ω-terminating. Next, we derive a result about fair derivations and ω-normal forms similar to the result in [DK89, DKP89, DKP90] by replacing the left linearity hypothesis with the retraction condition on the supporting algebra $T_\Sigma(V)_{/S}$. Then, we introduce and discuss some requirements on the infinite rewriting relation $\rightarrow_T$, which allow us to guarantee its ω-confluence.

Proposition 1 Let S be a term rewriting system of only unfolding rules. Then $\rightarrow_S$ is top-terminating.

Proof Straightforward. ♦

Proposition 2 The rewriting relation $\rightarrow_T = \rightarrow_{R/F} \cup \rightarrow_S$ on $T_\Sigma(V)_{/S}$ is top-terminating and ω-terminating.

Sketch of the proof Let us first show that $\rightarrow_T$ is top-terminating. Since R is the finite component of the infinite term rewriting system T, the hypothesis that R is finitely terminating implies that it is also top-terminating. Furthermore, from the definition of $\rightarrow_{R/F}$ (Section 2) it follows that $\rightarrow_{R/F}$ is also top-terminating and (ω-)terminating. On the other hand, $\rightarrow_S$ is top-terminating for Proposition 1. Thus, $\rightarrow_T$ is top-terminating. The fact that $\rightarrow_T$ is ω-terminating follows from top-termination by applying the same arguments as in Theorem 11 in [DKP89]. ♦

Proposition 3 Given the rewriting relation $\rightarrow_T$ and any term $t \in T_\Sigma(V)_{/S}$, then:

i.) if t admits an ω-normal form t', then it exists a fair derivation
$t = t_0 \rightarrow_T t_1 \rightarrow_T \ldots \rightarrow_T t_n \rightarrow_T \ldots$ with $\lim_{n\rightarrow\infty} t_n = t'$;

ii.) for any fair derivation $t = t_0 \rightarrow_T t_1 \rightarrow_T \ldots \rightarrow_T t_n \rightarrow_T \ldots$ with $\lim_{n\rightarrow\infty} t_n = t'$, t' is an ω-normal form of t.

Sketch of the proof

i.) The term t admits an ω-normal form t', hence (Definition 6) $t \rightarrow_T^{\omega} t'$ and t' cannot be reduced by $\rightarrow_{R/F}$, since such reduction does not preserve the limit. By contradiction, let us suppose that it does not exist a fair derivation which computes an ω-normal form of t. Let D be a derivation $t = t_0 \rightarrow_T t_1 \rightarrow_T \ldots \rightarrow_T t_n \rightarrow_T \ldots$ such that its limit $t' = \lim_{n\rightarrow\infty} t_n$ is an ω-normal form of t. Let us suppose that D is not fair. For Definition 7, there exists a "hanging" reduction by $\rightarrow_{R/F}$ along the derivation and it can be applied to the limit t' as well. This contradicts the hypothesis that t' is an ω-normal form.

ii.) Let D be a fair derivation $t = t_0 \rightarrow_T t_1 \rightarrow_T \ldots \rightarrow_T t_n \rightarrow_T \ldots$ with $\lim_{n\rightarrow\infty} t_n = t'$. By contradiction, let us suppose that t' is not an ω-normal form of t. Due to the fairness hypothesis we have only to consider the case in which t' can be reduced on an infinite redex by a non left linear rule in $\rightarrow_R$, whose application was never possible on any of the finite terms t_n along D. In fact, in order to be applied, such rules may require the equivalence of syntactically different subexpressions which represent the same infinite term. Thus, it could happen that a reduction by $\rightarrow_R$ is never detected on the finite terms in D, because it involves subexpressions which are semantically equivalent, but syntactically different. Such subexpressions become syntactically equivalent at the limit and the reduction can then be applied. Since the equivalence between subexpressions in $\rightarrow_R$ is checked modulo $=_F$, the described situation can never occur. ♦

Proposition 3 allows us to restrict our attention to fair derivations, as they compute ω-normal forms at the limit. Actually, there are cases in which it is possible to identify a subclass of fair derivations which have a peculiar structure.

Definition 11 (structured derivation) A derivation $t = t_0 \rightarrow_T t_1 \rightarrow_T \dots \rightarrow_T t_n \rightarrow_T \dots$ over $T_\Sigma(V)_{/S}$ is *structured* if there exists an index N such that, for all $n \geq N$, $t_n \rightarrow_S t_{n+1}$ and it never happens that $t_n \rightarrow_{R/F} t_{n+1}$ can be applied.

Thus, for any structured derivation it is possible to single out an index N which splits the infinite derivation into a finite subderivation of terms t_n $(n \leq N)$, in which $\rightarrow_{R/F} \cup \rightarrow_S$ is applied, and an infinite subderivation of terms t_n $(n > N)$, in which only $\rightarrow_S$ can be applied. Note that, in general, there is no guarantee that even a fair derivation is structured.

Example 2

The term rewriting system

R $\quad g(x, g(c,y)) \rightarrow g(x,y)$

S $\quad f(g(c, g(x,y))) \rightarrow g(c, g(x, f(g(c, g(x,y)))))$

allows the following fair derivation that is not structured since every rewriting step by $\rightarrow_S$ generates a reduction for $\rightarrow_{R/F}$:

D: $f(g(c, g(a,b))) \rightarrow_S g(c, g(a, f(g(c, g(a,b))))) \rightarrow_S g(c, g(a, g(c, g(a, f(g(c, g(a,b)))))))$
$\rightarrow_{R/F} g(c, g(a, g(a, f(g(c, g(a,b)))))) \rightarrow_S \dots$

Definition 12 (uniformity) A term rewriting system T is *uniform* if for any fair derivation D: $t \rightarrow_T t_1 \rightarrow_T \dots \rightarrow_T t_n \rightarrow_T \dots$ with $\lim_{n\to\infty} t_n = t'$, there exists a structured fair derivation D': $t \rightarrow_T t'_1 \rightarrow_T \dots \rightarrow_T t'_n \rightarrow_T \dots$ with $\lim_{n\to\infty} t'_n = t''$ and $t'' = t'$.

Our interest on uniform term rewriting systems is twofold. First, in order to show the ω-confluence of an infinite rewriting relation, it is possible to restrict only to the finite parts of the infinite derivations, thus retrieving all the results valid for finitely terminating rewriting relations, e.g. local confluence. Second, given a uniform term rewriting system, in general it is possible to determine a bound N on the number of rewriting steps of a fair derivation, which guarantees that a finite representation of the ω-normal form has been reached. In case of confluent uniform term rewriting systems, this means that it is possible to obtain a decision procedure for deciding the equivalence of two terms by computing their ω-normal forms.

In order to show the ω-confluence of $\rightarrow_T$ some additional requirements on the nature of R and S have to be stated.

Definition 13 Given $R = \{l_i \to r_i \mid 1 \le i \le n\}$ and $S = \{G_j \to H_j[G_j] \mid 1 \le j \le k\}$, then R and S are *independent* if and only if for $1 \le i \le n$ and $1 \le j \le k$, l_i and G_j do not overlap.

Example 3

In the following term rewriting system

R	$g(a) \to b$
S	$f(g(x)) \to f(f(g(x)))$

R and S are not independent. It is easy to verify that, for example, the term f(g(a)) admits an infinite number of fair derivations leading to different ω-normal forms.

Proposition 4 Let $R = \{l_i \to r_i \mid 1 \le i \le n\}$ and $S = \{G \to H[G]\}$ such that R is canonical, S consists of a single unfolding rule. If R and S are independent and $T = R/F \cup S$ is uniform, then T is ω-confluent.

Sketch of the proof Under the uniformity hypothesis we can restrict to fair structured derivations. Thus, in order to prove that $\to_{R/F} \cup \to_S$ is ω-confluent, we can restrict to show the ω-confluence of $\to_{R/F} \cup \to_S$ on the finite subderivations of fair structured derivations. In this case, $\to_{R/F} \cup \to_S$ can be treated as a finitely terminating rewriting relation and its ω-confluence can be shown by means of local confluence.

Thus, we have to show that whenever $t' \; {}_T\!\leftarrow t \to_T t''$ at the occurrences u and u' respectively, then there exists a term q such that $t' \overset{*}{\to}_T q \; {}_T\!\overset{*}{\leftarrow} t''$.

Since R is canonical and S is an unfolding rule, we have only to consider the cases in which t can be rewritten with both $\to_{R/F}$ and $\to_S$. Let us consider the two cases:

a) t/u and t/u' are disjoint redexes. Straightforward.

b) the redex t/u contains the redex t/u'.

Let us first consider the situation in which a redex for $\to_{R/F}$ contains a redex for $\to_S$. It follows from the definition of $\to_{R/F}$ ($=_F$ is induced by S) that t'' can be rewritten into t' by using $\to_{R/F}$. On the other hand, if a redex t/u for $\to_S$ contains a redex t/u' for $\to_{R/F}$, t/u is an istance $G\sigma$ of G for some substitution σ and, since R and S are independent, t/u' can only occurr if σ substitutes a variable x of G with an instance $l_i\sigma'$ for some $l_i \to r_i$ in R and substitution σ'. The following diagram shows how the confluence can be obtained:

$$\begin{array}{ccc} & t\,[G\sigma[x \leftarrow l_i\sigma']] & \\ & \downarrow_S \qquad\qquad \downarrow_{R/F} & \\ t\,[H[G]\sigma[x \leftarrow l_i\sigma']] & & t\,[G\sigma[x \leftarrow r_i\sigma']] \\ & \downarrow_{R/F} \qquad\qquad \downarrow_S & \\ & t\,[H[G]\sigma[x \leftarrow r_i\sigma']] & \end{array}$$

♦

Corollary Let $R = \{l_i \to r_i \mid 1 \le i \le n\}$ and $S = \{G_j \to H_j[G_j] \mid 1 \le j \le k\}$ such that R is canonical, S consists of unfolding rules whose left hand sides do not overlap. If R and S are independent and $T = R/F \cup S$ is uniform, then T is ω-confluent.

4. Final Remarks

In this section we briefly discuss some of the notions introduced in the paper, namely the retraction property of the supporting algebra and the independence requirement on the rewriting rules.

As regards the former, it is worth recalling that, when dealing with recursive expressions that are regular set of equations, results exist that allow to compute the unique canonical representative in the class of the terms with the same (tree) semantics, e.g. [CKV74]. This means that our notion of retraction actually permits coping with a reasonably interesting class of infinite rewriting systems.

On the other hand, independence is quite a strong condition on the syntactic nature of the rewriting rules. It is, anyhow, weaker than the "non overlapping" condition on the whole term rewriting system, which is till now required to guarantee ω-confluence in case of non terminating left linear system [DK89, DKP89, DKP90]. Nevertheless, we think that it can be possible to replace the independence requirement with a more semantic condition which deals with the infinite nature of the term. Future work concerns the definition of a notion of *preservation* between the components R and S of a term rewriting system, which guarantees that a reduction by R on a term denoting an infinite data structure cannot destroy its infinite nature. Thus, the term can only be rewritten into a term denoting another infinite data structure. Inspired by this notion, some syntactic conditions on term rewriting systems can still be determined in order to preserve S, that weaken the constraint on the left hand sides of the rules by permitting overlapping, but put some constraints on the right hand sides of the rules.

References

[BD89] Bachmair, L., Dershowitz, N. Completion for Rewriting modulo a Congruence, Theoretical Computer Science **67**, (1989), 173-201.

[CKV74] Courcelle, B., Kahn, G., Vuillemin, J. Algorithmes d'Equivalence et de Reduction a des Expressions Minimales dans une Classe d'Equations Recursives Simples, Proc.

Automata, Languages and Programming, 2nd Colloquium, Saarbrucken, LNCS **14**, (1974), 200-213.

[DIN90] De Nicola, R., Inverardi, P., Nesi, M. Using the Axiomatic Presentation of Behavioural Equivalences for Manipulating CCS Specifications, in Proc. Workshop on Automatic Verification Methods for Finite State Systems, LNCS **407** (1990), 54-67.

[DK89] Dershowitz, N., Kaplan, S. Rewrite, Rewrite, Rewrite, Rewrite, Rewrite..., Proc. 16th Annual ACM Symposium on POPL, Austin, Texas, (1989), 250-259.

[DKP89] Dershowitz, N., Kaplan, S., Plaisted, D.A. Infinite Normal Forms, Proc. 16th ICALP, Stresa, Italy, LNCS **372**, (1989), 249-262.

[DKP90] Dershowitz, N., Kaplan, S., Plaisted, D.A. Rewrite, Rewrite, Rewrite, Rewrite, Rewrite..., draft paper (1990).

[IN90a] Inverardi, P., Nesi, M. A Rewriting Strategy to Verify Observational Congruence, Information Processing Letters Vol. **35**, No. **4**, (August 1990), 191-199.

[IN90b] Inverardi, P., Nesi, M. On Deciding Observational Congruence of Finite-State CCS Expressions by Rewriting, I.E.I. Internal Report nº B4-10, (March 1990).

[Mil80] Milner, R. A Calculus of Communicating Systems, LNCS **92**, (1980).

[Mil89] Milner, R. A Complete Axiomatization for Observational Congruence of Finite-State Behaviours, Information and Computation **81**, (1989), 227-247.