

Modularity of UN^{\rightarrow} for left-linear term rewriting systems

M. Marchiori

Computer Science/Department of Software Technology

Report CS-R9433 May 1994

CWI is the National Research Institute for Mathematics and Computer Science. CWI is part of the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications.

SMC is sponsored by the Netherlands Organization for Scientific Research (NWO). CWI is a member of ERCIM, the European Research Consortium for Informatics and Mathematics.

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

Modularity of UN[→] for left-linear Term Rewriting Systems

Massimo Marchiori

CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
and
Department of Pure and Applied Mathematics
University of Padova
Via Belzoni 7, 35131 Padova, Italy
max@hilbert.math.unipd.it

Abstract

We give a positive solution to the open problem of whether the unicity of normal forms with respect to reduction (UN⁻) is a modular property for left-linear Term Rewriting Systems. Moreover, the 'pile and delete' technique employed allows for quite a short proof, and is of independent interest.

1991 Mathematics Subject Classification: 68Q42

1992 CR Categories: F.4.2

Keywords and Phrases: term rewriting system, left-linearity, modularity, m-collapsing.

1 Introduction

Modularity, that is the ability to solve a problem by solving its smaller subparts, is a fundamental topic in modern computer science. As far as Term Rewriting Systems (TRSs) are concerned, a property is called modular provided it is valid for two TRSs if and only if it holds for their disjoint sum. This area is nowadays a well established theory (see for instance [Klo90, Klo92, Mid90]). It is known of every important property whether it is modular or not; the only open problem in the field of unconditional ('pure') TRSs since 1989 ([Mid89, DJK91]) was the question of modularity for the uniqueness of normal form with respect to reduction (UN^{\rightarrow}) , the most basic among the normal form properties, in the case of left-linear TRSs. The two most important properties concerning uniqueness of normal forms are UN^{\rightarrow} and UN ([Klo92]): UN^{\rightarrow} states that every term has (at most) one normal form, whereas UN states that every term has (at most) one normal form modulo convertibility. It was proved by Middeldorp in [Mid89] that UN is a modular property. For UN^{\rightarrow} instead, this is no longer true. Indeed, take the two TRSs $\{a \rightarrow c, a \rightarrow e, b \rightarrow d, b \rightarrow e, e \rightarrow e\}$ and $\{F(X, X) \rightarrow A\}$: the term F(a, b) has two distinct normal forms, namely F(c, d) and A.

As said, in the case of left-linear TRSs the modularity of UN⁻ was still unknown: in this paper we give a proof to this open question, showing that indeed UN⁻ is a modular property. The main idea is not to try to analyze the complex behaviour that a general reduction in the direct sum of two TRSs can have, but instead to modify it to get a simpler one.

First, a suitable definition of modular marking of a term is introduced; this naturally leads

Report CS-R9433 ISSN 0169-118X to the formulation of the key concept of modular collapsing (m-collapsing), that will prove to be essential. Indeed, it is shown that, provided only that the TRS is left-linear, failure of UN⁻ cannot occur without m-collapsings.

Finally, using a technique called 'pile and delete', every possible counterexample to the modularity of UN[¬] is translated into one without m-collapsings, thus obtaining a contradiction. This technique, besides allowing for a rather concise proof, turned out to be important on its own, since thanks to it also a new easy and short proof of the modularity of completeness (see [TKB89]) has been given in [Mar94] (even more: see Subsection 4.1). Moreover, since its application does not require the full power of UN[¬] but the weaker property of consistency with respect to reduction (CON[¬]), stating that a term cannot be rewritten to two different variables, the proof here given also yields the result that CON[¬] is modular for left-linear Term Rewriting Systems.

The paper is organized as follows: after giving the necessary preliminaries in Section 2, Section 3 introduces the concepts of modular marking and modular collapsing, showing their relevance in the study of UN^{\rightarrow} ; finally, Section 4 proves the main theorem stating modularity of UN^{\rightarrow} for left-linear TRSs by means of the 'pile and delete' technique, and shows that, via the same proof, CON^{\rightarrow} is modular for left-linear TRSs as well.

2 Preliminaries

We assume familiarity with the basic notions regarding Term Rewriting Systems: the notation used is essentially the one in [Klo92] and [Mid90].

Contexts will be denoted as usual with \square and with square brackets $(C[\cdot \cdot \cdot])$. Throughout the paper we will indicate with \mathcal{A} and \mathcal{B} the two TRSs to operate on, their corresponding sets of function symbols with $\mathcal{F}_{\mathcal{A}}, \mathcal{F}_{\mathcal{B}}$, the set of variables as \mathcal{V} , and the set of terms built from some set of function symbols \mathcal{F} and \mathcal{V} as $\mathcal{T}(\mathcal{F})$.

When not otherwise specified, all symbols and notions not having a TRS label are to be intended operating on the disjoint sum $\mathcal{A} \oplus \mathcal{B}$. For better readability, we will talk of function symbols belonging to \mathcal{A} and \mathcal{B} like white and black functions, indicating the first ones with upper case functions, and the second ones with lower case. Variables, instead, have no colour.

Definition 2.1 The *root* symbol of a term $t \in \mathcal{T}(\mathcal{F}_{\mathcal{A}} \cup \mathcal{F}_{\mathcal{B}})$ is f if $t \equiv f(t_1, ..., t_n)$, and t itself otherwise.

Let $t \equiv C[t_1, ..., t_n] \in \mathcal{T}(\mathcal{F}_1 \cup \mathcal{F}_2)$ and $C \not\equiv \square$; we write $t \equiv C[t_1, ..., t_n]$ if C[...] is an $\mathcal{F}_{\mathcal{A}}$ -context and each of the t_i has $root(t_i) \in \mathcal{F}_{\mathcal{B}}$, or vice versa (exchanging \mathcal{A} and \mathcal{B}). The topmost homogeneous part (briefly top) of a term $C[t_1, ..., t_n]$ is the context $C[\cdot \cdot \cdot]$.

Definition 2.2 The rank of a term $t \in \mathcal{T}(\mathcal{F}_{\mathcal{A}} \cup \mathcal{F}_{\mathcal{B}})$ is 1 if $t \in \mathcal{T}(\mathcal{F}_{\mathcal{A}})$ or $t \in \mathcal{T}(\mathcal{F}_{\mathcal{B}})$, and $\max_{i=1}^n \{rank(t_i)\} + 1$ if $t \equiv C[[t_1, \ldots, t_n]]$ (n > 0).

The following well known lemma will be implicitly used in the sequel:

Lemma 2.3 $s \rightarrow t \Rightarrow rank(s) \geq rank(t)$

Proof. Clear.

Definition 2.4 The multiset S(t) of the special subterms of a term t is

1.
$$S(t) = \begin{cases} \{t\} & \text{if } t \in (\mathcal{T}(\mathcal{F}_{\mathcal{A}}) \cup \mathcal{T}(\mathcal{F}_{\mathcal{B}})) \setminus \mathcal{V} \\ \emptyset & \text{if } t \in \mathcal{V} \end{cases}$$
2.
$$S(t) = \bigcup_{i=1}^{n} S(t_i) \cup \{t\} \text{ if } t \equiv C[t_1, \dots, t_n] \quad (n > 0)$$

Note that this definition is slightly different from the usual ones in the literature (for example in [Mid90]), since here variables are not considered to be special subterms.

If $t \equiv C[t_1, ..., t_n]$, the t_i are called the *principal* special subterms of t. Furthermore, a reduction step of a term t is called *outer* if the rewrite rule isn't applied in the principal special subterms of t.

A (strict) partial order on the special subterms of a term can be naturally given defining $t_1 > t_2$ iff t_2 is a proper special subterm of t_1 (that is $t_2 \in \mathcal{S}(t_1), t_1 \neq t_2$).

The following proposition will reveal useful:

Proposition 2.5 If A and B are left-linear, then rewrite rules that have the possibility to act outer on a special subterm t are exactly those that have the possibility to act on its top.

Proof. Let $t = C[t_1, ..., t_n]$: since $t_1, ..., t_n$ have a root belonging to the other TRS (with respect to C), they are matched by variables from any rewrite rule applicable to C, and for the left-linearity assumption these variables are independent each other. \Box Note that left-linearity is essential for this proposition.

3 Marking and Collapsing

To be able to describe the special subterms of a given term throughout a reduction, it is natural to develop a concept of (modular) marking. A first, naïve approach of modular marking for a term is to take an assignment from the multiset of its special subterms to a (fixed) set of markers. So, for instance, given the term F(f(G,a),H), we could mark $F(\Box,H)$ to m_1 , $f(\Box,a)$ to m_2 , G to m_3 . Then reductions steps, as usual, should preserve the markers. However, this simple definition presents a problem, since for one case there is ambiguity: when a collapsing rule makes an *inner* top vanish. In this case, we have the following situation (see Figure 1):

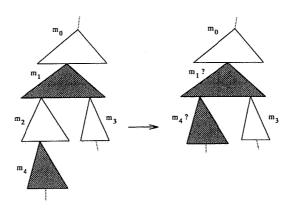


Figure 1: Naïve modular marking.

and we have a conflict between m_1 and m_4 .

This situation is dealt with by defining a modular marking for a term to be an assignment from the multiset of its special subterms to sets of markers, and taking in the ambiguous case just described the union of the marker sets of the two special subterms involved.

Thus, the previous example would give (singletons like $\{m_3\}$ are written simply m_3):

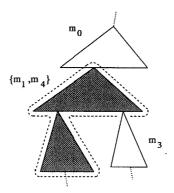


Figure 2: Correct modular marking.

When this situation occurs, we say that the special subterm m_4 has been absorbed by m_1 , and the special subterm m_2 has had a modular collapsing (briefly m-collapsing). This last concept will reveal to be crucial in the study of the UN \rightarrow property (Theorem 3.4).

When dealing with reductions t woheadrightarrow t' we will always assume, in order to distinguish all the special subterms, that the initial modular marking of t is injective and maps special subterms to singletons.

Inside a reduction a notion of descendant for every special subterm can be defined: in a reduction a special subterm is a descendant of another if the set of markers of the former contains the set of markers of the latter. Note, en passant, that due to the presence of duplicating rules, there may be more than one descendant, or even none (due to erasing rules).

Summing up, there is a complete description of what happens to the descendants of a special subterm:

Proposition 3.1 A special subterm, when a reduction step is applied, can only: i) be erased ii) m-collapse iii) be preserved (i.e. have descendants).

Proof. Obvious.

3.1 Left-linearity and UN[→]

When the left-linearity and UN→ properties are introduced, m-collapsings enjoy some remarkable properties. First of all, they behave in a 'deterministic' way, in the following sense:

Proposition 3.2 Let A be left-linear and UN^{\rightarrow} , and $t = C[[t_1, ..., t_n]]$ a top white special subterm. Then, if t m-collapses into t_i $(1 \le i \le n)$ via a white reduction (i.e. using only rules from A), the index i is unique.

Proof. Since \mathcal{A} is left-linear, by Proposition 2.5 the white reduction depends only on the top of t. Hence, if we take instead of $t = C[[t_1, \ldots, t_n]]$ a term $t' = C[[X_1, \ldots, X_n]]$ (with X_1, \ldots, X_n new fresh variables), then every previous white reduction that m-collapsed t to t_i can be repeated on t' to reduce it to X_i , and if the index i were not unique t' could be reduced to different normal forms, contradicting the fact \mathcal{A} is UN^{\rightarrow} .

Moreover, the concept of m-collapsing reveals to be crucial in the study of UN^{\rightarrow} modularity for the following reason:

Definition 3.3 A UN^{\rightarrow} counterexample (briefly counterexample) is a pair (d_1, d_2) , where $d_1: s \rightarrow s_1 \rightarrow \ldots \rightarrow s_{k_1}$ and $d_2: s \rightarrow t_1 \rightarrow \ldots \rightarrow t_{k_2}$ are reductions starting from the same term s (called the start) and ending in two normal forms $s_{k_1} \not\equiv t_{k_2}$ (called the ends).

Theorem 3.4 If A and B are left-linear and UN^{\rightarrow} , then there is no counterexample without m-collapsings.

Proof. Take a reduction without m-collapsings ending in a normal form. Every rule acts on the top of a a well specified special subterm, and this top cannot change since no m-collapsing is present. Moreover, by Proposition 2.5, the application of these rules depends only on the top itself. So for every top of a special subterm a separate reduction is performed, that must eventually lead to a unique top for the UN^{\rightarrow} property, and hence the resulting normal form is unique as well.

4 Pile and Delete

The pile and delete technique here employed allows, once given a term and some reductions that normalize it, to transform the given term (and correspondingly the reductions too) in such a way to preserve the set of normal forms previously obtained, but this time with reductions in a nice form, that is without m-collapsings.

Proposition 4.1 If A and B are left-linear and UN^{\rightarrow} , every counterexample can be translated into a counterexample without m-collapsings.

Proof. If the counterexample is already without m-collapsings, the assertion is trivially satisfied. So, suppose it is not. Select a special subterm of the start of the counterexample that has rank minimal amongst the ones that m-collapse in the counterexample itself: say $t = \tau[t_1, \ldots, t_n]$. This special subterm cannot stay (have descendants) till an end of the counterexample without being absorbed. Indeed, suppose it is so, and t reaches a normal form n. Because of its rank minimality, t must m-collapse by Proposition 3.2 into a fixed principal subterm, namely t_i . So, substituting (in t) t_i with a new fresh variable X, we can obtain, by Proposition 2.5, a reduction from this new term t' to the normal form X which is without m-collapsings.

On the other hand, t also reduces to the normal form n different from X via a reduction without m-collapsings (again, by the minimality assumption) and so, by Proposition 2.5, unregarding of what is in t_i . Therefore also t' reduces to n via the same reduction, giving a counterexample without m-collapsings, in contradiction with Theorem 3.4.

The fact that t alone cannot reach the ends does not mean that its top, τ , is not needed at all in the counterexample: it may be needed, via absorption, from other white tops of

 (\prec) -greater special subterms in the counterexample. All of these special subterms $\bar{r}_1, \ldots, \bar{r}_\ell$ are descendants of some special subterms of the start r_1, \ldots, r_k $(k \leq \ell)$.

We can so try to perform 'in advance' these absorptions, modifying directly the start of the counterexample, using the following 'pile and delete' technique.

First, we 'pile' $\tau[t_1, \ldots, t_{i-1}, \square, t_{i+1}, \ldots, t_n]$ just below the tops of the r_1, \ldots, r_k . That is to say if $r_i = r_i[s_1, \ldots, s_v]$ and t is in s_j (viz. $t \prec s_j$), then r_i is replaced with

$$r_i[s_1,\ldots,s_{j-1},\tau[t_1,\ldots,t_{i-1},s_j,t_{i+1},\ldots,t_n],s_{j+1},\ldots,s_v]$$

The situation is illustrated in Figure 3.

r₁

after:

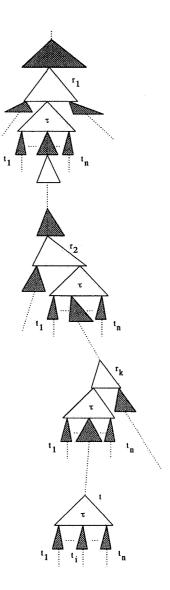


Figure 3: The 'pile' process.

Intuitively, the top of t isn't really needed any more, since we have already inserted copies of it where needed for absorption, and it has been proved earlier that t alone cannot stay till an end of the counterexample: therefore we 'delete' it replacing t by t_i (see Figure 4).

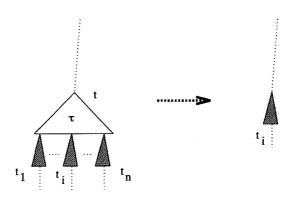


Figure 4: The 'delete' process.

Now it has to be shown that the original counterexample can still be mimicked using this revised start term; this can be done because we can get rid of the piled τ , when not needed, using the original reduction from the counterexample that m-collapsed it $(t \rightarrow t_i)$:

- By minimality of t, the only effect of the rules acting on the descendants of t but not on the descendants of t_i was to m-collapse t into a descendant of t_i (if this is not the case, then it means that the descendant of t must be erased), and so they can be dropped since we already replaced t with t_i .
- When a descendant of t was absorbed by, say, \bar{r}_q , we have piled to its ancestor r_p (and so to its descendant \bar{r}_q) in that place $\tau[t_1, \ldots, t_{i-1}, \square, t_{i+1}, \ldots, t_n]$, whereas the old descendant of t is now the corresponding descendant of t_i . So it only remains to reduce the piled $\tau[t_1, \ldots, t_{i-1}, \square, t_{i+1}, \ldots, t_n]$ as previously in the counterexample to obtain exactly the same situation as before, and the new counterexample can proceed in the mimicking (see Figure 5). Note how these postponed reductions produce no m-collapsings.
- We inserted $\tau[t_1, \ldots, t_{i-1}, \square, t_{i+1}, \ldots, t_n]$ below all the r_1, \ldots, r_k , but actually descendants of t may be absorbed in the initial counterexample only by part of the descendants of these special subterms. However, we can get rid of these superfluous occurrences of material acting, as hinted previously, with the rules that in the initial counterexample made $\tau[t_1, \ldots, t_{i-1}, \square, t_{i+1}, \ldots, t_n]$ collapse into \square : they are applied to all of these extra descendants when the piled material is not needed any more. This means that these 'deleting sequences' must be applied
 - i) when in the sequel of the original reduction the descendant of an r_p will not absorb a descendant of τ any more, or
 - ii) when the descendant of an r_p absorbs another descendant of an r_q (Figure 6).

Again, it is immediate to see that these deleting sequences produce no m-collapsings.

This way we have obtained a new counterexample with a different start but the same ends as the initial one. Once again, note that left-linearity, via Proposition 2.5, was essential to be able to mimic the old counterexample.

Now consider the number of special subterms of the start with a descendant that m-collapses in the counterexample itself: this new counterexample obtained via the pile and delete technique has this number diminished by one with respect to the initial counterexample; indeed, t is no more present, and as remarked no new m-collapsings are introduced modifying the original reductions.

So, repeating this 'pile and delete' process leads, ultimately, to a counterexample without m-collapsings.

Example 4.2 Consider the two following (left-linear and UN→) TRSs

$$\mathcal{A} = \left\{ egin{array}{ll} F(X)
ightarrow G(X,X) \ G(L(X,Y),Z)
ightarrow Y \ H(X,Y)
ightarrow L(X,Y) \end{array}
ight. \quad \mathcal{B} = \left\{ egin{array}{ll} f(X)
ightarrow X \ g(f(X))
ightarrow a \ g(X)
ightarrow g(X) \end{array}
ight.$$

and the reduction (unary functions like f(A) are for short written fA from now on)

$$gFfH(A, fA) \rightarrow gFfH(A, fB) \rightarrow gG(fH(A, fB), fH(A, fB)) \rightarrow$$

 $gG(fH(A, fB), H(A, fB)) \rightarrow gG(fH(A, fB), H(A, B)) \rightarrow$
 $gG(H(A, fB), H(A, B)) \rightarrow gG(L(A, fB), H(A, B)) \rightarrow gfB \rightarrow a$

The special subterm of the starting term with minimal rank among the ones that m-collapse in this reduction is fA. Thus, after the pile and delete process we get

$$gfFfH(A,A) \rightarrow gfFfH(A,B) \rightarrow gfG(fH(A,B),fH(A,B)) \rightarrow gfG(fH(A,B),H(A,B)) \rightarrow gfG(H(A,B),H(A,B)) \rightarrow gfG(L(A,B),H(A,B)) \rightarrow gfB \rightarrow a$$

Now the minimal special subterm of the starting term that m-collapses is fH(A,A); the corresponding reduction after the pile and delete is

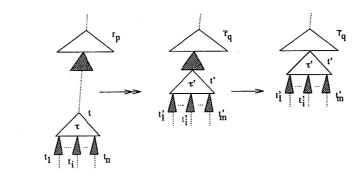
$$gfFH(A,A) \rightarrow gfFH(A,B) \rightarrow gfG(H(A,B),H(A,B)) \rightarrow gfG(L(A,B),H(A,B)) \rightarrow gfB \rightarrow a$$

and this reduction is without m-collapsings.

Theorem 4.3 UN^{\rightarrow} is a modular property for left-linear TRSs.

Proof. One verse is obvious. On the other hand, if \mathcal{A} and \mathcal{B} are UN^{\rightarrow} but $\mathcal{A} \oplus \mathcal{B}$ is not, then it has a counterexample that can be translated into a counterexample without m-collapsings by the above Proposition 4.1, contradicting Theorem 3.4.

Old reduction:



 $New\ reduction:$

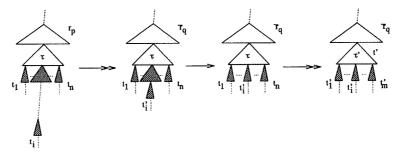
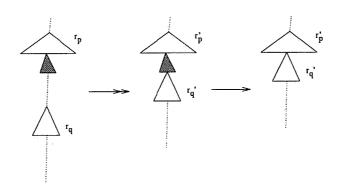


Figure 5: One case of mimicking.

Old reduction:



New reduction:

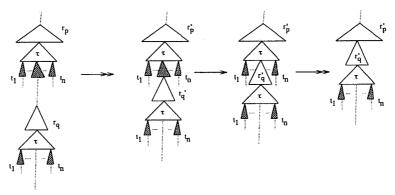


Figure 6: An application of a deleting sequence.

4.1 Weakening UN→

The 'pile and delete' technique does not need the full power of UN^{\rightarrow} , but it can be applied under the weaker assumption of consistency with respect to reduction (briefly CON^{\rightarrow}), that is satisfied if a term cannot be rewritten to two different variables. This is true since the 'pile and delete' technique essentially relies upon Proposition 3.2, that still holds if CON^{\rightarrow} is required in place of UN^{\rightarrow} . Hence, if we replace the definition of UN^{\rightarrow} -counterexample with the corresponding definition of CON^{\rightarrow} -counterexample (where the ends are required to be variables), exactly the same proof here used for the modularity of UN^{\rightarrow} shows that

Theorem 4.4 CON^{\rightarrow} is a modular property for left-linear TRSs.

This result, together with the modularity of UN^{\rightarrow} for left-linear TRSs, fills the gap in the parallelism between the pairs (UN, UN^{\rightarrow}) and (CON, CON^{\rightarrow}) (a TRS is consistent, CON for short, if different variables cannot be equal modulo convertibility). Indeed, this parallelism is present in all the other cases, since:

- UN \Rightarrow UN \rightarrow and CON \Rightarrow CON \rightarrow (straightforward).
- UN ≠ UN[→] (see [Mid89]) and CON ≠ CON[→] (take the TRS $\{f(X) \to X, f(X) \to a\}$ that is CON[→] but $X \leftarrow f(X) \to a \leftarrow f(Y) \to Y$).
- UN is modular unlike UN⁻ ([Mid89]) and CON is modular unlike CON⁻ (modularity of CON has been proved by Manfred Schmidt-Schauß in [SS89], whereas to see that CON⁻ is not modular take the two TRSs $\{f(X) \to X, f(X) \to a\}$ and $\{F(X, X, Y) \to Y, F(X, Y, Y) \to X\}$ that are CON⁻ but in their disjoint sum $X \leftarrow f(X) \leftarrow F(f(X), a, a) \leftarrow F(f(X), f(Y), f(Z)) \twoheadrightarrow F(a, a, f(Z)) \to f(Z) \to Z$.

Furthermore, using the fact CON→ is modular for left-linear TRSs and that CON→ suffices to apply the 'pile and delete' technique, in [Mar94], using the same proof method employed here, it has been not only given a new easy and short proof of the modularity of completeness for left-linear TRSs (cf. [TKB89]), but also proved the modularity of termination for left-linear and consistent with respect to reduction TRSs.

Acknowledgments

I would like to heartly thank Jan Willem Klop: besides giving many profitable remarks regarding the writing style, his love for simplicity was a continuous stimulus to get shorter and better proofs from the original version of this paper. Thanks are due to Aart Middeldorp for careful reading this paper and for having implicitly suggested me to try to weaken the UN⁻ property. Thanks also to Enno Ohlebush.

References

[DJK91] N. Dershowitz, J. Jouannaud, and J.W. Klop. Open problems in rewriting. In R.V. Book, editor, Proceedings of the Fourth International Conference on Rewriting Techniques and Applications, volume 488 of LNCS, pages 445-456. Springer-Verlag, 1991.

- [Klo90] J.W. Klop. Term rewriting systems. Technical Report CS-R9073, CWI, Amsterdam, December 1990.
- [Klo92] J.W. Klop. Term rewriting systems. In S. Abramsky, Dov M. Gabbay, and T.S.E. Maibaum, editors, Handbook of Logic in Computer Science, volume 2, chapter 1, pages 1–116. Clarendon Press, Oxford, 1992.
- [Mar94] M. Marchiori. Modularity of completeness revisited. Technical Report 6, Department of Pure and Applied Mathematics, University of Padova, Italy, 17 May 1994.
- [Mid89] A. Middeldorp. Modular aspects of properties of term rewriting systems related to normal forms. In N. Dershowitz, editor, *Proceedings of the Third International Conference on Rewriting Techniques and Applications*, volume 355 of *LNCS*, pages 263–277. Springer-Verlag, April 1989.
- [Mid90] A. Middeldorp. Modular Properties of Term Rewriting Systems. PhD thesis, Vrije Universiteit, Amsterdam, November 1990.
- [SS89] M. Schmidt-Schauß. Unification in a combination of arbitrary disjoint equational theories. J. Symbolic Computation, 8(1,2):51-99, 1989.
- [TKB89] Y. Toyama, J.W. Klop, and H.P. Barendregt. Termination for the direct sum of left-linear term rewriting systems. In N. Dershowitz, editor, Proceedings of the 3rd International Conference on Rewriting Techniques and Applications, volume 355 of LNCS, pages 477-491, Chapel Hill, April 1989. Springer-Verlag. Extended version: Report CS-R8923, CWI, Amsterdam.