Computing Constructor Forms with Non Terminating Rewrite Programs
Isabelle Gnaedig, Hélène Kirchner

To cite this version:

HAL Id: inria-00112083
https://inria.hal.science/inria-00112083
Submitted on 7 Nov 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Abstract
In the context of the study of rule-based programming, we focus in this paper on the property of C-reducibility, expressing that every term reduces to a constructor term on at least one of its rewriting derivations. This property implies completeness of function definitions, and enables to stop evaluations of a program on a constructor form, even if the program is not terminating. We propose an inductive procedure proving C-reducibility of rewriting. The rewriting relation on ground terms is simulated through an abstraction mechanism and narrowing. The induction hypothesis allows assuming that terms smaller than the starting terms rewrite into a constructor term. The existence of the induction ordering is checked during the proof process, by ensuring satisfiability of ordering constraints. The proof is constructive, in the sense that the branch leading to a constructor term can be computed from the proof trees establishing C-reducibility for every term.

Categories and Subject Descriptors F.3.1 [LOGICS AND MEANINGS OF PROGRAMS]: Specifying and Verifying and Reasoning about Programs—Logics of programs, Mechanical verification, Specification techniques; F.4.2 [MATHEMATICAL LOGIC AND FORMAL LANGUAGES]: Grammars and Other Rewriting Systems; F.4.3 [MATHEMATICAL LOGIC AND FORMAL LANGUAGES]: Formal Languages—Algebraic language theory; I.1.3 [SYMBOLIC AND ALGEBRAIC MANIPULATION]: Languages and Systems—Evaluation strategies, Substitution mechanisms; I.2.3 [ARTIFICIAL INTELLIGENCE]: Deduction and Theorem Proving—Deduction, Inference engines, Mathematical induction; D.3.1 [PROGRAMMING LANGUAGES]: Formal Definitions and Theory; D.2.4 [SOFTWARE ENGINEERING]: Software/Program Verification—Correctness proofs, Formal methods, Validation

General Terms Algorithms, Languages, Verification

Keywords Sufficient Completeness, Constructor, Abstraction, Narrowing, Ordering Constraint, Termination

1. Introducing the problem
Sufficient completeness plays an important role in algebraic specifications, as well as in rewriting-based programs. In both contexts, terms are built on operators, among them we can distinguish constructors. Constructors are basic symbols, allowing us to describe the values of computations. The other symbols, called defined symbols, represent functions defined on these values. One is naturally led to ask for a specification or a program, whether the functions are sufficiently defined to warrant that to every expression or term corresponds a value.

From the point of view of specifications, where properties are described by equations, sufficient completeness ensures that every term is equivalent to a term built on constructors, called constructor term or constructor form. It allows inductive proofs, in particular by consistency methods [7]. Proof assistants like Coq or PVS include decision or semi-decision procedures based on rewrite rules and rely on complete definitions of functions.

From the point of view of programming, sufficient completeness ensures that a program produces a completely computed form for every data. Studying the property in this context is pertinent since rule-based systems have recently gained considerable interest with the development of efficient compilers. Now, systems like ASF+SDF [25], Maude [5, 12], Cafe-OBJ [17], ELAN [1], or TOM [29], are used for various applications like constraint solving, protocol verification, modeling of biological or chemical systems, and more.

Proving sufficient completeness is undecidable in general. It has already been widely studied, for example in [21, 30, 26, 23, 6, 22, 28, 2, 3], but most of the time, the proposed approaches for proving the property need restrictions like termination and confluence.

The property is strongly related to ground reducibility, which expresses that every ground instance of a term is reducible. Indeed, it is equivalent to ground reducibility of all patterns \( f(x_1, \ldots, x_m) \) built on a defined symbol \( f \), provided the rewrite system (RS in short) is terminating, confluent, and the normal form of a constructor term is again a constructor term [22]. Under these conditions, techniques developed for proving ground reducibility hold for sufficient completeness as well [31, 24, 22, 27, 4, 8].

In this paper, we address the problem of sufficient completeness from the programming point of view, and we go beyond the previous usual restrictions.

We observe the case where a program or a RS can be neither confluent, nor terminating, and we study its evaluations. We do not suppose other restrictions used in the domain as the constructor preserving property, or the absence of relation between constructors. The question is to know whether at least one evaluation of a given data gives a completely evaluated result; in other words, whether for every ground term, there exists a rewriting chain that eventually reaches a constructor term, even if the chains do
not converge to a single term, and are infinite. We call this property C-reducibility.

Since several years, we have been studying properties of rewriting in the context of programming, where a rewrite system interpreter or compiler uses a specific strategy for applying rewrite rules, most frequently an innermost strategy that consists in rewriting always at the lowest possible positions, and corresponds to a call-by-value evaluation of functions. We observe rewriting in considering its properties from an inductive point of view. We already have proposed proof procedures for the termination property, in the specific case of rewriting strategies, respectively for the innermost [20], the outermost [14] and local strategies [13]. We also have given a constructive proof technique for weak innermost termination, enabling to reach the terminating branches [15]. Here, C-reducibility is tackled with a similar inductive approach.

It is not based on analyzing the patterns issued from the RS, as do most of existing methods for sufficient completeness, but on observing the rewriting process starting from any term to see whether it is eventually transformed into a constructor term. For that, like for our previous procedures for termination, we reason by induction on the property itself, assuming that the terms \( t \) smaller than the starting terms for an induction ordering are C-reducible. This approach needs a noetherian ordering on terms, used in the induction principle. In contrast to classical induction proofs where the ordering is given, we do not need to define it a priori. We only have to check its existence by ensuring satisfiability of ordering constraints incrementally set along the proof.

By alternatively applying an abstraction mechanism and well-covered narrowing steps, we generate proof trees representing at least one rewriting derivation for every ground term. The proof is constructive, in the sense that the rewriting branch leading to a constructor term can be computed from the proof trees establishing C-reducibility for every term, thus avoiding costly breadth-first computations to reach the constructor form. Although our method differs from previous approaches, we naturally encounter basic notions already used for proving sufficient completeness, as unification, used in [28], or covering properties used in many works in the domain, for example in [26, 28]. In [3], pattern trees are also developed, but under the assumption that the RS is terminating and ground confluent.

Before we detail our procedure, we link C-reducibility to close properties like strong C-reducibility, expressing the existence of a constructor form on every rewriting chain, and their variants in the case of innermost rewriting. Sufficient completeness and ground reducibility are also considered in this comparison. In particular, C-reducibility directly implies sufficient completeness. In addition, we justify that ground reducibility just requires weak termination (expressing that every term has at least one terminating rewriting derivation) to imply sufficient completeness, thus weakening the condition of the well-known theorem of [22]. We also note that as the covering property we require is stronger than ground reducibility, this property directly implies sufficient completeness if we suppose weak termination. Thus, with respect to the sufficient completeness property, the proof technique we propose is interesting:

- obviously when the RS is not even weakly terminating,
- when it is terminating, or just weakly terminating, but we have no technique to prove it. The weak termination proof is difficult to handle in general, and to our knowledge, our proof procedure [15] is the only one existing today. On the following small example, weakly terminating, and which is a classical definition of the booleans, enriched with a rule expressing the double application of \( \text{not} \) on \( \text{and} \), and deliberately oriented in the divergent direction, it fails:

\[
\begin{align*}
\text{and}(1, x) & \rightarrow x \\
\text{and}(0, x) & \rightarrow 0 \\
\text{or}(1, x) & \rightarrow 1 \\
\text{or}(0, x) & \rightarrow x \\
\text{and}(1, x) & \rightarrow \text{not} (\text{not}(\text{and}(1, x))) \\
\text{not}(1) & \rightarrow 0 \\
\text{not}(0) & \rightarrow 1 \\
\text{not}(\text{and}(x, y)) & \rightarrow \text{or} (\text{not}(x), \text{not}(y)).
\end{align*}
\]

- as an alternative of the termination proof, as on the following more realistic example of computation of quotient, which is innermost terminating:

\[
\begin{align*}
\text{quot}(0, s(y), s(z)) & \rightarrow 0 \\
\text{quot}(s(x), s(y), z) & \rightarrow \text{quot}(x, y, z) \\
\text{quot}(x, 0, s(z)) & \rightarrow \text{quot}(x, s(z), s(z)) \\
\text{quot}(x, y, 0) & \rightarrow \text{error} \\
\text{quot}(\text{error}, y, z) & \rightarrow \text{error} \\
\text{quot}(x, \text{error}, z) & \rightarrow \text{error} \\
\text{quot}(x, y, \text{error}) & \rightarrow \text{error}.
\end{align*}
\]

In addition, if the RS has no rule with a constructor left-hand side, our proof technique establishes weak termination at the same time. This is the case for the two examples above.

In Section 2, the background is presented. In Section 3, we link known notions of completeness to C-reducibility. Section 4 introduces the basic concepts of the inductive proof mechanism. In Section 5, our procedure is defined with inference rules and a strategy to apply them. In Section 6, we explain how to extract a rewriting chain leading to a constructor term, from the proof of C-reducibility of terms.

2. The background

We assume that the reader is familiar with the basic definitions and notations of term rewriting given for instance in [11]. \( T(\mathcal{F},X) \) is the set of terms built from a given set \( \mathcal{F} \) of function symbols having an arity \( ar = n \in \mathbb{N} \), and a set \( X \) of variables denoted \( x, y, \ldots \). In this paper, we only consider finite sets of symbols. \( T(\mathcal{F}) \) is the set of ground terms (without variables). Symbols of arity 0 are called constants. Positions in a term are represented as sequences of integers. The empty sequence \( \epsilon \) denotes the top position. The notation \( t_{1..p} \) stands for the subterm of \( t \) at position \( p \). To emphasize that \( u \) contains subterms \( t_{1..p} \), we write \( u[t_1..t_p] \). We also write \( u[t_1..t_p] \).

A substitution is an assignment from \( X \) to \( T(\mathcal{F},X) \), written \( \sigma \in (x \mapsto t) \ldots (y \mapsto u) \). It uniquely extends to an endomorphism of \( T(\mathcal{F},X) \). We identify a substitution \( \sigma = (x \mapsto t) \ldots (y \mapsto u) \) with the finite conjunction of equations \( (x = t) \land \cdots \land (y = u) \). The result of applying \( \sigma \) to a term \( t \in T(\mathcal{F},X) \) is written \( \sigma(t) \) or \( t\sigma \). The domain of \( \sigma \), denoted \( \text{Dom}(\sigma) \), is the finite subset of \( X \) such that \( \sigma x \neq x \). The range of \( \sigma \), denoted \( \text{Range}(\sigma) \), is the set \( \bigcup_{x \in \text{Dom}(\sigma)} \text{Var}(\sigma x) \).

A ground substitution or instantiation is an assignment from \( X \) to \( T(\mathcal{F}) \). The composition of substitutions \( \sigma_1 \) followed by \( \sigma_2 \) is denoted \( \sigma_2 \sigma_1 \). Given two substitutions \( \sigma_1 \)
and $\sigma_2$, we write $\sigma_1 \leq \sigma_2$ iff $\exists \theta$ such that $\sigma_2 = \theta \sigma_1$. We say that $\sigma_1$ is a generalization of $\sigma_2$.

For $F$, we declare a set $C$ of constructors and a set $D$ of defined symbols. The terms of $T(C)$ are called constructor terms, or more briefly $C$-terms (or $C$-forms). Given a set $R$ of rewrite rules or term rewriting system on $T(F,X)$, the rewriting relation induced by $R$ is denoted by $\rightarrow_R$ ($\rightarrow$ if there is no ambiguity on $R$). We denote $s \rightarrow^{p,l-r,\sigma} t$ (where either $p$ or $l \rightarrow r$ or $\sigma$ may be omitted) if $s$ rewrites into $t$ at position $p$ with the rule $l \rightarrow r$ and the substitution $\sigma$, i.e. $s = s[\sigma]_p$ and $t = s[\sigma r]_p$. The reflexive transitive closure of the rewriting relation induced by $R$ is denoted by $\rightarrow_R^*$. Given a term $t$, we call normal form of $t$ any irreducible term $u$, if it exists, such that $t \rightarrow_R^* u$. If a term $t$ rewrites to a $C$-form, we write this $C$-form $tl_C$, and say that it is a $C$-reduced form for $R$.

An ordering $\succ$ on $T(F,X)$ is said to be noetherian iff there is no infinite decreasing chain for this ordering. It is monotone iff for every pair of terms $t,t'$ of $T(F,X)$, for every context $f(\ldots,t_t,\ldots)$, $t\succ t'$ implies $f(\ldots,t'_t,\ldots) \succ f(\ldots,t'_t,\ldots)$. It has the subterm property iff for every $t$ of $T(F,X)$, $f(\ldots,t_1,\ldots,t_n) \succ t$. It has the constructor subterm property iff for every $g(t_1,\ldots,t_n)$ of $T(F,X)$ with $g \in C$, we have $g(t_1,\ldots,t_n) \succ t_1,\ldots,t_n$. Note that if $\succ$ is monotone and has the subterm property, as $F$ is finite, then it is noetherian [10]. If, in addition, $\succ$ is stable by substitution (for every substitution $\sigma$, every pair of terms $t,t' \in T(F,X)$, $t \succ t'$ implies $t[\sigma] \succ t'[\sigma]$), then it is called a simplification ordering. A precedence is an ordering on $F$, denoted $\preceq_F$.

A RS $R$ (innermost) terminates iff every (innermost) rewriting derivation (or derivation chain) for $R$ is finite. It weakly terminates if at least one derivation chain starting from every term is finite. It is confluent iff for every term $t$ such that $u \rightarrow^* t$ and $t \rightarrow^* v$, there exists $t'$ such that $u \rightarrow^* t'$ and $v \rightarrow^* t'$. It is constructor preserving iff for each rule $l \rightarrow r$ of $R$, whenever $l \in T(C)$, then $r \in T(C)$.

3. C-reducibility and related properties

Before we compare them, let us now formally define the properties presented in the introduction.

**Definition 1.** Let $R = \{ l_i \rightarrow r_i, i \in [1..m] \}$ be a RS on $T(F,X)$. The equational theory associated to $R$ is the equational theory induced by the set of equations $E = \{ l_i = r_i, i \in [1..m] \}$. It is also noted $E$.

**Definition 2.** Let $E$ be an equational theory on $T(F,X)$, where $F = C \cup D$. $E$ is sufficiently complete (with respect to $D$) if for every term $t \in T(F)$, there is a term $u \in T(C)$ such that $t \equiv_E u$.

**Definition 3.** Let $R$ be a RS on $T(F,X)$. A term $t \in T(F,X)$ is ground reducible if every ground instance of $t$ is reducible with $R$.

**Definition 4.** Let $R$ be a RS on $T(F,X)$, where $F = C \cup D$. A ground term is (innermost) (strongly) C-reducible iff it reduces into a $C$-term on at least one of its (resp. all) (innermost) rewriting derivations. The RS $R$ is said to be (innermost) (strongly) C-reducing iff every term $t$ of $T(F)$ is (innermost) (strongly) C-reducible.

Let us make some remarks on this property. If $R$ does not have rules whose left-hand sides are $C$-terms, the reached $C$-forms are normal forms.

If there are such rules in $R$, which is a case of relations between constructors, the $C$-forms may be reducible. Then, either $R$ is constructor-preserving and the next terms in the rewriting derivations are still $C$-terms. Or $R$ is not constructor-preserving and terms with defined symbols can be introduced after the $C$-terms in the derivations.

In non terminating cases, C-reducibility allows us to introduce a weak pseudo-termination notion, expressing that the evaluation of every data stops on the first constructor encountered form, ensured to exist on at least one rewriting branch.

Let us now observe the following diagram, summarizing the links between the different properties we consider. “Ground reducibility” stands for ground reducibility of all patterns $f(x_1,\ldots,x_m), f \in D$; “C-reducibility” stands for C-reducibility of all terms of $T(F)$. Arrows represent implications between properties; they are labeled by the necessary conditions on the RS required by the implications. All links are quite obvious, except those between ground reducibility and sufficient completeness, given by Theorem 7 in [22].

From the diagram, we immediately set the following theorem, which weakens the terminating condition of the “if” part of Theorem 7 in [22].

**Theorem 1.** Let $R$ be an innermost terminating RS on $T(F,X)$, where $F = C \cup D$. If every term $f(x_1,\ldots,x_m), f \in
\( \mathcal{D}, x_1, \ldots, x_m \in \mathcal{X} \) is ground reducible, then the equational theory \( \mathcal{E} \) associated to \( \mathcal{R} \) is sufficiently complete.

This result can be verified in analyzing the proof of Theorem 7 in [22], where the termination hypothesis can be indeed weakened into innermost termination, and even into weak termination.

Note that sufficient completeness does not imply \( \mathcal{C} \)-reducibility in every case: for a given term, a \( \mathcal{C} \)-form reachable by \( \mathcal{E} \) can be unreachably by the rewriting relation, as illustrates the following terminating but not confluent example, where \( a, b, c \) are constants and \( c \) is the only constructor form:

\[
\begin{align*}
  a & \rightarrow b \\
  a & \rightarrow c.
\end{align*}
\]

4. Inductively proving \( \mathcal{C} \)-reducibility

For proving that a term \( t \) of \( \mathcal{T}(\mathcal{F}) \) is \( \mathcal{C} \)-reducible, we proceed by induction on \( \mathcal{T}(\mathcal{F}) \) with a nonempty ordering \( \succ \), assuming the property for any \( t' \) such that \( t \succ t' \). To warrant non emptiness of \( \mathcal{T}(\mathcal{F}) \) and \( \mathcal{T}(\mathcal{C}) \), we assume that \( \mathcal{C} \) contains at least a constant. The main intuition is to observe rewriting derivations starting from a ground term \( t \in \mathcal{T}(\mathcal{F}) \) which is any instance of a term \( g(x_1, \ldots, x_m) \in \mathcal{T}(\mathcal{F}, \mathcal{X}) \), for some defined function symbol \( g \in \mathcal{D} \), and variables \( x_1, \ldots, x_m \). Proving the \( \mathcal{C} \)-reducibility property on ground terms amounts to prove that at least one derivation leads to a constructor term.

Rewriting derivations are simulated, using a lifting mechanism, by a proof tree developed from a ground term \( g(x_1, \ldots, x_m) \) on \( \mathcal{T}(\mathcal{F}, \mathcal{X}) \), for every \( g \in \mathcal{D} \), by alternatively using two main concepts, namely narrowing and abstraction. More precisely, narrowing schematizes the rewriting possibilities of terms. Abstraction simulates the reduction of subterms in the derivations until these subterms become \( \mathcal{C} \)-terms. It expresses the application of the induction hypothesis on these subterms.

The schematization of ground rewriting derivations is achieved through constraints. The nodes of the developed proof trees are composed of a current term of \( \mathcal{T}(\mathcal{F}, \mathcal{X}) \), and a set of ground substitutions represented by a constraint progressively built along the successive abstraction and narrowing steps. Each node in an abstract tree schematizes a set of ground terms: all ground instances of the current term, that are solutions of the constraint.

The constraint is in fact composed of two kinds of formulas: ordering constraints, set to warrant the validity of the inductive steps, and abstraction constraints combined to narrowing substitutions, which effectively define the relevant sets of ground terms.

As said previously, we consider any term of \( \mathcal{T}(\mathcal{F}) \) as a ground instance of a term \( t \) of \( \mathcal{T}(\mathcal{F}, \mathcal{X}) \) occurring in a proof tree issued from a reference term \( t_{ref} \).

- the resulting term \( u = t[X_j]_{\{i_1, \ldots, i_p\}} \) (where \( i_1, \ldots, i_p \) are the abstraction positions in \( t \)) is narrowed in all possible ways into terms \( v \), according to the possible instances of the \( X_j \). This corresponds to rewriting ground instances of \( u \) (characterized by the constraint associated to \( u \)) with all possible rewrite rules.

- This technique is inspired from the one we have proposed for proving innermost termination. But in that case, as abstracted subterms were representing innermost normal forms, alternating abstraction and innermost narrowing steps allowed us to simulate all innermost rewriting derivations. Here, abstracted subterms do not generally represent normal forms, and we use standard narrowing, so the process only simulates some possible rewriting derivations, for the standard rewriting relation (without strategy) among which we try to prove the existence of a \( \mathcal{C} \)-term.

- However, we have to ensure that every ground term represented by \( u \) is \( \mathcal{C} \)-reducible, and so we have to prove the existence of at least one branch leading to a \( \mathcal{C} \)-term for every ground instance of \( u \).

- To exhibit such a branch for every ground instance of \( u \) containing at least one defined symbol, the narrowing steps have to cover (in a sense we will define later) all possible ground instances of \( u \).

- Then \( \mathcal{C} \)-reducibility of the ground instances of \( t \) is reduced to \( \mathcal{C} \)-reducibility of the ground instances of the terms \( v \). Now, if \( \theta_{ref} \succ \theta v \) for every ground substitution \( \theta \) that is a solution of the constraint associated to \( v \), by the induction hypothesis, \( \theta v \) is supposed to be \( \mathcal{C} \)-reducible. Else, the process is iterated on \( v \), until we get a term \( t' \) such that either \( \theta_{ref} \succ \theta \theta' \) or \( \theta \theta' \) is a \( \mathcal{C} \)-term.

- We now introduce some concepts to formalize and automate this mechanism.

4.1 Ordering constraints and abstraction

The induction ordering \( \succ \) is constrained along the proof by imposing constraints between terms that must be comparable, each time the induction hypothesis is used in the abstraction mechanism. As we are working with a lifting mechanism on the proof trees with terms of \( \mathcal{T}(\mathcal{F}, \mathcal{X}) \), we directly work with an ordering \( \succ \succ \) on \( \mathcal{T}(\mathcal{F}, \mathcal{X}) \) such that \( t_{ref} \succ \succ u \) induces \( \theta_{ref} \succ \theta u \), for every \( \theta \) solution of the constraint associated to \( u \).

- So inequalities of the form \( t_{ref} \succ u_1, \ldots, u_m \) are accumulated, which will be called ordering constraints. Any ordering \( \succ \succ \) on \( \mathcal{T}(\mathcal{F}, \mathcal{X}) \) satisfying them and which is stable by substitution fulfills the previous requirements on ground terms. The ordering \( \succ \succ \), defined on \( \mathcal{T}(\mathcal{F}, \mathcal{X}) \), can then be seen as an extension of the induction ordering \( \succ \), defined on \( \mathcal{T}(\mathcal{F}) \).

- For convenience, the ordering \( \succ \succ \) will also be written \( \succ \).

- It is important to note that, for establishing the inductive termination proof, it is sufficient to decide whether there exists such an ordering.

**Definition 5.** An ordering constraint is a pair of terms of \( \mathcal{T}(\mathcal{F}, \mathcal{X}) \) noted \( (t \succ t') \). It is said to be satisfiable if there exists an ordering \( \succ \), such that for every instantiation \( \theta \) whose domain contains \( \text{Var}(t) \cup \text{Var}(t') \), we have \( \theta \succ \theta' \). We say that \( \succ \) satisfies \( (t \succ t') \).

A conjunction \( C \) of ordering constraints is satisfiable if there exists an ordering satisfying all conjuncts. The empty conjunction, always satisfied, is denoted by \( \top \).

Satisfiability of a constraint \( C \) of this form is undecidable. But a sufficient condition for an ordering \( \succ \) to satisfy \( C \)
is that $\triangleright$ is stable by substitution and $t \triangleright t'$ for every constant $t \triangleright t'$ of $C$.

The inductive termination proof on ground terms by application of Theorem 2 requires that the non-erasing induction ordering has the constructor subterm property. Simplification orderings fulfill such conditions. So in practice, it is sufficient to find a simplification ordering $\triangleright$ such that $t \triangleright t'$ for any constraint $t \triangleright t'$ of $C$.

Solving ordering constraints in finding simplification orderings is a well-known problem in rewriting. The simplest and automatatable way to proceed is to test simple existing orderings like the subterm ordering, the Recursive Path Ordering, or the Lexicographic Path Ordering. This is often sufficient for the constraints considered here. If these simple orderings are not powerful enough, automatic solvers like Cine2 [9] can provide adequate polynomial orderings.

Other constraints are introduced by the abstraction mechanism. As said previously, to abstract a term $t$ at positions $i_1, \ldots, i_p$, where the $t_{j}$, $j \in \{i_1, \ldots, i_p\}$ are supposed to have a reduced $C$-form $t_{ij}$, we replace the $t_{ij}$ by abstraction variables $X_j$ representing respectively one of their possible $C$-forms. Let us define these special variables more formally.

**Definition 6.** Let $N$ be a set of new variables disjoint from $X$. Symbols of $N$ are called $C$-variables (or constructor variables). Substitutions and instantiations are extended to $T(F,X \cup N)$ in the following way. Let $X \in N$; for any substitution $\sigma$ (resp. instantiation $\theta$) such that $X \in \text{Dom}(\sigma)$, $\sigma X \in T(C,N)$ (resp. $\theta X \in T(C)$) and then $\text{Var}(\sigma X) \subseteq N$.

**Definition 7 (abstraction).** The term $t\{t_{i_1}\}_{i_1, \ldots, i_p}$ is said to be abstracted into the term $u$ (called abstraction of $t$) at positions $i_1, \ldots, i_p$ iff $u = t[X_1]_{i_1, \ldots, i_p}$, where the $X_j$, $j \in \{i_1, \ldots, i_p\}$ are fresh distinct $C$-variables.

$C$-reducibility on $T(F)$ is proved by reasoning on terms with abstraction variables, i.e. on terms of $T(F,X \cup N)$. Ordering constraints are extended to pairs of terms of $T(F,X \cup N)$. When subterms $t_{i}$ are abstracted by $X_i$, we state constraints on abstraction variables, called abstraction constraints, to express that their instances can only be $C$-terms, rewritten from the corresponding instances of $t_{i}$. Initially, they are of the form $t_{i} = X_{i}$ where $t \in T(F,X \cup N)$, and $X_{i} \in N$, but we will see later on how they will be combined with the substitutions used for the narrowing process.

### 4.2 Narrowing

After abstraction of the current term $t$ into the term $u$, if $u$ contains defined symbols, we test whether the ground instances of $u$ are reducible, according to the possible values of the instances of the $X_{i}$. This is achieved by narrowing $u$ in all possible ways. We recall the definition of narrowing.

**Definition 8 (narrowing).** A term $t \in T(F,X \cup N)$ narrows into a term $t' \in T(F,X \cup N)$ at the non-variable position $p$, using the rule $l \rightarrow r \in R$ with the substitution $\sigma$ which is written $t \rightarrow^{p,l \rightarrow r,\beta \sigma} t'$ iff

$$\sigma(l) = \sigma(t[p]) \text{ and } t' = \sigma(t[r])$$

where $\sigma$ is the most general unifier of $t$ at position $p$ and $l$.

A few remarks can be made on the choice of variables and on the domain of the substitutions generated during the proof process. It is always assumed that there is no variable in common between the rule and the term, i.e. that $\text{Var}(t) \cap \text{Var}(t) = \emptyset$. This requirement of disjoint variables is easily fulfilled by an appropriate renaming of variables in the rules when narrowing is performed. Observe that for the most general unifier $\sigma$ used in the above definition, $\text{Dom}(\sigma) \subseteq \text{Var}(t) \cup \text{Var}(t)$ and we can choose $\text{Ran}(\sigma) \cap (\text{Var}(t) \cup \text{Var}(t)) = \emptyset$, thus introducing in the range of $\sigma$ only fresh variables. Moreover, narrowing is only performed on terms $t$ of $T(F,N)$, since an abstracting step is first applied on the reference terms, of the form $g(x_1, \ldots, x_m)$, replacing $x_1, \ldots, x_m$ by $X_1, \ldots, X_m \in N$. Then from Definition 6 we infer that for the most general unifiers $\sigma$ produced during the proof process, all variables of $\text{Ran}(\sigma)$ are $C$-variables.

Notice also that in our process, we are interested in the narrowing substitution applied to the current term $t$, but not in its definition on the variables of the left-hand side of the rule. So the narrowing substitutions we consider are restricted to the variables of the narrowed term $t$.

As said before, we only narrow the term $u$ if all narrowing possibilities represent at least one rewriting step for all possible ground instances of $u$.

Else, the proof process stops with failure. But we cannot conclude that $t$ is not $C$-reducible, since, although all ground instances of $u$ are not reducible at that step, they may be reduced to a $C$-form by other rewriting derivations, not observed by the process, as shows the following example. Let $F = \{f, a, b, c\}$, $C = \{b, c\}$, $R = \{a \rightarrow b, f(f(b)) \rightarrow c, b \rightarrow f(b), f(c) \rightarrow c\}$. Abstracting the pattern $f(x)$ gives $f(X)$, where $X$ only represents $C$-forms. Then, $f(X)$ is only narrowed into $c$ for $X = c$. The narrowing substitution $X = f(b)$ is not allowed since $f(b) \notin T(C,N)$. However, the other ground instance of $f(X)$, which is $f(b)$, is also $C$-reducible, but the rewriting chain $f(b) \rightarrow f(f(b)) \rightarrow c$ is not represented by the abstract-narrow-based process.

**Definition 9.** A set of substitutions $\Sigma$, is said to cover a term $u \in T(F,X \cup N)$, or to be $u$-covering iff for any ground instance $\theta u$ of $u$, $\exists \sigma \in \Sigma$ such that $\text{Dom}(\sigma) \supseteq \text{Var}(u)$, and $\theta u = \sigma u$ for some ground substitution $\mu$.

The following obvious proposition expresses that Definition 9 provides a sufficient condition for the previous requirement on narrowing.

**Proposition 1.** Let $R$ be a RS, $u$ a term of $T(F,N)$, and $\Sigma$ the set of narrowing substitutions of $u$ for $R$. If $\Sigma$ is $u$-covering, then every ground instance $\alpha u$ of $u$ is such that $\alpha u \rightarrow_{p,l \rightarrow r,\beta \sigma} t'$, for some ground substitution $\beta$, and we have $u \rightarrow_{p,l \rightarrow r,\beta \sigma} v$ for some $v$ of $T(F,N)$, $\beta \sigma = \alpha$ on any variable set $\gamma \supseteq \text{Var}(u) \cup \text{Dom}(\alpha)$, and $t' = \beta v$.

As a first narrowing step is applied on the patterns $g(x_1, \ldots, x_m)$, $g \in D$, issued from abstraction of the reference patterns $g(x_1, \ldots, x_m)$, the development of each proof tree at least requires for the narrowing substitutions of $g(x_1, \ldots, x_m)$ to cover $g(x_1, \ldots, x_m)$. The following two propositions allow us to show that the covering property of the patterns $g(x_1, \ldots, x_m)$, $g \in D$ is stronger than usual ground reducibility (i.e. of the patterns $g(x_1, \ldots, x_m)$). The first one is an obvious consequence of Proposition 1.

**Proposition 2.** Let $R$ be a RS, $g \in D$, $X_1, \ldots, X_m \in N$. Let $\Sigma$ be the set of narrowing substitutions of $g(x_1, \ldots, x_m)$ with $R$. If $\Sigma$ covers $g(x_1, \ldots, x_m)$, then $g(x_1, \ldots, x_m)$ is ground reducible.
The converse is not true, as shows the following example.
Let \( F = \{ f, 0, 1 \} \), \( C = \{ 0, 1 \} \), and \( R = \{ f(0) \rightarrow 0, 1 \rightarrow 0 \} \).
The term \( f(X) \) is ground reducible since \( f(0) \) and \( f(1) \) are reducible. However, the set \( \Sigma \) of narrowing substitutions of \( f(X) \), equal to \( \{ \sigma = \langle X = 0 \rangle \} \) is not covering for \( f(X) : f(1) \) is not a ground instance of \( f(0) \).

**Proposition 3.** Let \( R \) be a RS on \( T(F, X) \), \( X_1, \ldots, X_m \in N \), \( x_1, \ldots, x_m \in X \). Then \( f(X_1, \ldots, X_m) \) is ground reducible for every \( f \in D \), \( \sigma \) iff \( f(x_1, \ldots, x_m) \) is ground reducible for every \( f \in D \).

Given a term \( u \), a sufficient condition for a substitution set \( \Sigma \) to be \( u \)-covering can be established as follows, provided the variables of the considered term \( u \) are \( C \)-variables, whose ground instantiations can only be constructor terms. Let \( P \) be the set of constructor patterns \( \{ f(Y_1, \ldots, Y_m) \mid f \in C, Y_1, \ldots, Y_m \in N, x_1, \ldots, x_m \in X \} \).

**Proposition 3.** Let \( R \) be a RS on \( T(F, X) \), \( X_1, \ldots, X_m \in N \), \( x_1, \ldots, x_m \in X \). Then \( f(X_1, \ldots, X_m) \) is ground reducible for every \( f \in D \), \( \sigma \) iff \( f(x_1, \ldots, x_m) \) is ground reducible for every \( f \in D \).

4.3 Cumulating constraints

Abstraction constraints have to be combined with the narrowing substitutions to characterize the ground terms schematized by the current term \( t \) in the proof tree. Indeed, a narrowing branch on the current term \( u \) with narrowing substitution \( \sigma \) represents a rewriting branch for any ground instance of \( \sigma u \).

In addition, \( \sigma \) has to satisfy the constraints on variables of \( u \), already set in \( A \). So \( \sigma \), considered as the narrowing constraint attached to the narrowing branch, is added to \( A \). This leads to the introduction of abstraction constraint formulas.

**Definition 10.** An abstraction constrained formula \( (ACF \text{ in short}) \) is a formula \( \land_{i = 1}^{m} \langle t_{1i} = t' \rangle \land \land_{i = 1}^{m} \langle x_i = u_i \rangle \), where \( x_i \in X \cup N \), \( t_1, \ldots, t_m \in T(F, X \cup N) \).

**Definition 11.** An abstraction constrained formula \( A = \land_{i = 1}^{m} \langle t_{1i} = t' \rangle \land \land_{i = 1}^{m} \langle x_j = u_j \rangle \) is satisfiable iff there exists at least one instantiation \( \theta \) such that \( \land_{i = 1}^{m} \langle \theta t_{1i} = \theta t'_i \rangle \land \land_{i = 1}^{m} \langle \theta x_i = \theta u_i \rangle \). The instantiation \( \theta \) is then said to satisfy the ACF \( A \) and is called solution of \( A \).

For a better readability on examples, we can propagate \( \sigma \) into \( A \) (by applying \( \sigma \) to \( A \)), thus getting instantiated abstraction constraints of the form \( t_{1i} = t'_i \) from initial abstraction constraints of the form \( t_{1i} = X_i \).

An ACF is attached to each term \( u \) in the proof tree; the ground substitutions solutions of \( A \) define the instances of the current term \( u \) in the proof tree, for which we are observing \( C \)-reducibility. When \( A \) has no solution, the current node of the proof tree represents no ground term. Such nodes are then irrelevant for the proof. Detecting and suppressing them during a narrowing step allows us to control the narrowing mechanism. So we have the choice between generating only the relevant nodes of the proof tree, by testing satisfiability of \( A \) at each step, or stopping the proof on a branch on an irrelevant node, by testing unsatisfiability of \( A \).

Checking satisfiability of \( A \) in general undecided, but it is often easy in practice to exhibit an instantiation satisfying it. Automatable sufficient conditions are also under study.

Unsatisfiability of \( A \) is also undecidable in general, but automatable simple sufficient conditions can be used, as testing whether \( A \) contains equalities \( t_{1i} = u \), where \( u \) contains at least a symbol of \( D \).

In the following, we present the procedure dealing with satisfiability of \( A \).

5. Inference rules for the inductive proof

We are now ready to describe the inference rules defining our mechanism. They transform a set \( T \) of \( 3 \)-tuples \( (U, A, C) \) where \( U = \{ t \} \) or \( \emptyset \), \( t \) is the current term whose \( C \)-reducibility has to be proved, \( A \) is an abstraction constraint formula, \( C \) is a conjunction of ordering constraints.

- **The first rule** abstracts the current term \( t \) at given positions \( t_1, \ldots, t_p \). The constraints \( t_{ref} > t_1, \ldots, t_p \) are set, allowing us to suppose, by induction, the existence of \( C \)-forms for \( t_1, \ldots, t_p \). Then, \( t_{1}, \ldots, t_{p} \) are abstracted into abstraction variables \( X_1, \ldots, X_p \). The abstraction constraint \( t_{1i} \mid C = X_1, \ldots, t_{pi} \mid C = X_p \) is added to the ACF \( A \). We call that rule **Abstract**.

The abstraction positions are chosen so that the abstraction mechanism captures the greatest possible number of rewriting steps: then we abstract all greatest possible subterms of \( f(t_1, \ldots, t_n) \). More concretely, we try to abstract \( t_1, \ldots, t_m \) and, for each \( t_i = g(t'_i, \ldots, t'_{k_i}) \) that cannot be abstracted, we try to abstract \( t_1, \ldots, t_{i-1}, t_{i+1}, \ldots, t_m \) and so on. In the worst case, we are driven to abstract leaves of the term, which are either variables, or constants.

Note also that it is not useful to abstract subterms if they are in \( T(C, N) \). Indeed, by Definition 6, every ground instance of such subterms is already a \( C \)-term.

- **The second rule** narrows the resulting term \( u \), if it is not a term of \( T(C, N) \), in all possible ways in one step, with all possible rewrite rules of the rewrite system \( R \), and all possible substitutions \( \sigma_1, \ldots, \sigma_m \), into terms \( v_1, \ldots, v_k \), according to Definition 8, if \( \{ \sigma_1, \ldots, \sigma_m \} \) is \( u \)-covering. This step is a branching step, creating as many states as narrowing possibilities. The substitution \( \sigma \) is integrated to \( A \). This is the **Narrow** rule. As \( \{ \sigma_1, \ldots, \sigma_m \} \) is \( u \)-covering, the narrowing step effectively simulates at least one rewriting step for each possible ground instance of \( u \).

- **We finally have a Stop** rule halting the proof process on the current branch of the proof tree, when the ground instances of the current term can be stated as \( C \)-reducible. This happens when the whole current term \( u \) can be abstracted, i.e. when the induction hypothesis applies on it, or when \( u \in T(C, N) \).

As said before, ordering constraints have to be satisfied by a noetherian ordering having the constructor subterm property. Any simplification ordering holds. We can make further assumptions on the ordering to enable constraints to be satisfied. In particular, for a Recursive Path Ordering or Lexicographic Path Ordering whose precedence \( \succ \) is such that \( f \succ f, c, \forall f \in D, \forall c \in C \), we have \( t \succ u \) for \( t \in T(F, X) \) containing at least a symbol of \( D \) and \( u \in T(C) \).

Exploiting equalities in \( A \) can also help to solve ordering constraints: \( t_{1i} = X_i \) means that \( \theta X_i \) is a \( C \)-form obtained by rewriting \( \theta t_i \), for every ground substitution \( \theta \). As we do not abstract terms of \( T(C, N) \), if \( t_i \neq x_i \in X \), \( \theta t_i \) contains at least a symbol of \( D \) and then \( \theta t_i > \theta X_i \) for a RPO or an
LPO defined as above. As \( t_{ref} \) also contains a symbol of \( D \), we also have \( \Theta_{ref} \supset \theta X \) for \( X \in \mathcal{N} \).

Before giving the inference rules, let us note that the inductive reasoning can be completed in the following way. When the induction hypothesis cannot be applied to a term \( u \), it may be possible to prove \( C \)-reducibility of every ground instance of \( u \) by another way. Let \( C \rightarrow \text{RED}(u) \) be a predicate that is true iff every ground instance of \( u \) is \( C \)-reducible. In the first and third previous steps of the inductive reasoning, we then associate the alternative predicate \( C \rightarrow \text{RED}(u) \) to the condition \( t \supset u \). It is true in particular if \( u \in T(C, \mathcal{N}) \), as said in the presentation of the Stop rule.

The rules are given in Table 1. They use a reference term \( t_{ref} = g(x_1, \ldots, x_m) \), where \( x_1, \ldots, x_m \in \mathcal{X} \) and \( g \in D \) (if \( g \) is a constant, then \( t_{ref} = g \)). To prove \( C \)-reducibility for every term \( t \in T(F) \) we proceed as follows: for each defined symbol \( g \in D \), we apply the rules using the reference term \( t_{ref} = g(x_1, \ldots, x_m) \) on the initial set of 3-tuples \( \left\{ \left( \left( t_{ref} = g(x_1, \ldots, x_m) \right), (\top, \top) \right) \right\} \), with a specific strategy \( S \).

The strategy \( S \) repeats the following steps: first, apply Abstract, and then try Stop. Then try all possible applications of Narrow. Narrow, then Stop again.

Note that if \( A \) is satisfiable, the transformed forms of \( A \) by Abstract and Stop are also satisfiable. Moreover, the first application of Abstract generates \( A = (\wedge, x_1 \perp = X_1) \), always satisfied by the constructor constant supposed to exist in \( C \). Thus with the strategy \( S \), it is useless to prove satisfiability of \( A \) in the rules Abstract and Stop.

The process may not terminate if there is an infinite number of applications of Abstract and Narrow on the same branch of a derivation tree. Nothing can be said in that case. It stops if no inference rule applies anymore.

A branch of the derivation tree is said to be successful if it is ended by an application of Stop, i.e. if its final state is of the form \((\emptyset, A, C)\).

Thus, the inductive \( C \)-reducibility proof is successful if there is at least one successful branch in the proof tree, corresponding to each possible ground term. Let us develop this point.

In fact, branching, produced by Narrow, can generate different states with narrowing substitutions \( \sigma_1, \ldots, \sigma_n \). These substitutions can be compared. For \( \sigma_i \) and \( \sigma_j \), three situations may occur: \( \sigma_i \) is strictly less general than \( \sigma_j \), which is noted \( \sigma_i \prec \sigma_j \), (or \( \sigma_j \) is strictly less general than \( \sigma_i \)), \( \sigma_i \) and \( \sigma_j \) are equal up to a renaming, or else \( \sigma_i \) and \( \sigma_j \) are incomparable.

States corresponding to substitutions that are more general than other ones then represent a set of ground instances that contains the other ones. So, for proving \( C \)-reducibility for all ground instances at a branching point, it is sufficient to prove \( C \)-reducibility only for the “most general states”.

A branching node in a proof tree can only be a state, on which the Narrow rule applies. Let \( \Sigma \) be the set of narrowing substitutions (possibly with different rewrite rules) at a given branching node. Let \( \Sigma_0 \) be the reduced set from \( \Sigma \) such that \( \sigma \in \Sigma_0 \) iff \( \sigma \in \Sigma \) and \( \beta \sigma' \in \Sigma \) such that \( \sigma \prec \sigma' \) on \( (\text{Dom}(\sigma) \setminus \text{Var}(l)) \cup (\text{Dom}(\sigma') \setminus \text{Var}(l')) \), where \( l \) and \( l' \) are the left-hand sides of rules respectively used to produce the narrowing substitutions \( \sigma \) and \( \sigma' \). The set \( \Sigma_0 \) may yet contain equivalent (equal up to a renaming) substitutions which are marked as such. So for any two substitutions in \( \Sigma_0 \), either they are equivalent, or they are incomparable.

A proof tree is weakly successful if at each branching node:

- for each class of equivalent substitutions, there exists at least one weakly successful subtree corresponding to a substitution in this class,
- all subtrees corresponding to incomparable substitutions are weakly successful,
- a tree reduced to the state \((\emptyset, A, C)\) is weakly successful.

So the strategy \( S \) can be optimized as follows: at each branching point of a proof tree, with set of substitutions \( \Sigma \), we only develop the subtrees corresponding to \( \Sigma_0 \). Moreover, given two subtrees corresponding to equivalent substitutions, as soon as one of them is weakly successful, the other one is cut.

We write \( \text{SUCCESS}(g, \succ) \) if the proof tree obtained by application on \( \left\{ \left( g(x_1, \ldots, x_m) \right), (\top, \top) \right\} \), with the strategy \( S \), of the inference rules whose conditions are satisfied by an ordering \( \succ \), is weakly successful.

**Theorem 2.** Let \( R \) be a RS on a set \( \mathcal{F} \) of symbols. If there exists a noetherian ordering \( \succ \) having the constructor subterm property, such that for each defined symbol \( g \), we have \( \text{SUCCESS}(g, \succ) \), then \( R \) is \( C \)-reducing.

Recall the important point that the ordering \( \succ \) has to be the same for all \( g(x_1, \ldots, x_m) \in D \).

The subtree cut process is similar to the one defined for weak termination, for which a formal description with a complete set of inference rules is given in [16].

Let us come back to the first example of the introduction.

**Example:** We prove that the RS

\[
\begin{align*}
\text{and}(1, x) & \rightarrow x \\
\text{and}(0, x) & \rightarrow 0 \\
\text{or}(1, x) & \rightarrow 1 \\
\text{or}(0, x) & \rightarrow x \\
\text{and}(1, x) & \rightarrow \text{not}(\text{not}(\text{and}(1, x))) \\
\text{not}(1) & \rightarrow 0 \\
\text{not}(0) & \rightarrow 1 \\
\text{not}(\text{and}(x, y)) & \rightarrow \text{or}(\text{not}(x), \text{not}(y))
\end{align*}
\]

on \( T(F) \), with \( F = \{ \text{and} : 2, \text{or} : 2, \text{not} : 1, 1, 0, 0, 0 \} \), and \( C = \{ 0, 1 \} \), is \( C \)-reducing.

Applying the rules on \( \text{not}(x) \), we get:

\[
\begin{align*}
\text{not}(x_1) & \rightarrow A = \top, \ C = \top \\
\text{Abstract} & \\
\text{not}(X_1) & \rightarrow A = (x_1 \perp = X_1) \rightarrow \top \\
& \rightarrow C = (\text{not}(x_1) > x_1)
\end{align*}
\]

**Narrow**

0

\[
\begin{align*}
\sigma_1 & = (X_1 = 0) \\
A & = (x_1 \perp = 0) \\
C & = (\text{not}(x_1) > x_1)
\end{align*}
\]

\[
\begin{align*}
\sigma_2 & = (X_1 = 1) \\
A & = (x_1 \perp = 1) \\
C & = (\text{not}(x_1) > x_1)
\end{align*}
\]

**Stop(twice)**

\[
\begin{align*}
\emptyset & \rightarrow A = (x_1 \perp = 0) \\
& \rightarrow C = (\text{not}(x_1) > x_1)
\end{align*}
\]

\[
\begin{align*}
\emptyset & \rightarrow A = (x_1 \perp = 1) \\
& \rightarrow C = (\text{not}(x_1) > x_1)
\end{align*}
\]

**Narrow** would also apply with \( \sigma_1 = (X_1 = \text{and}(X_2, X_3)) \), but \( \sigma_2 \) is not licit since \( \text{and}(X_2, X_3) \notin T(C, N) \).
Moreover $\Sigma = \{\sigma_1, \sigma_2, \sigma_3\}$ is not($X_1$)-covering since for every possible ground instance $\theta X_1$ of $X_1$ (0 or 1), there exists $\sigma$ in $\Sigma$ such that $\theta$ not($X_1$) = $\mu \sigma$ not($X_1$) for some $\mu$, which is here equal to Id.

Stop applies (twice) since 0 and 1 are terms of $T(c, N)$.

The ordering constraints are satisfied by any term ordering having the subterm property, for example by a RPO with any precedence.

Now, considering $or(x_1, x_2)$, we get:

$\text{or}(x_1, x_2) \quad A = T \quad C = T$

Abstract

$\text{or}(X_1, X_2) \quad A = (x_1|_c = X_1 \land x_2|_c = X_2)
C = (\text{or}(x_1, x_2) > x_1, x_2)$

Narrow

$X_2
\sigma_1 = (X_1 = 0)
A = (x_1|_c = 0 \land x_2|_c = X_2)
C = (\text{or}(X_1, X_2) > x_1, x_2)$

$\sigma_2 = (X_1 = 1)
A = (x_1|_c = 1 \land x_2|_c = X_2)
C = (\text{or}(X_1, X_2) > x_1, x_2)$

Stop (twice)

$\emptyset
A = (x_1|_c = 0 \land x_2|_c = X_2)
C = (\text{or}(X_1, X_2) > x_1, x_2)$

$\emptyset
A = (x_1|_c = 1 \land x_2|_c = X_2)
C = (\text{or}(X_1, X_2) > x_1, x_2)$

$\Sigma = \{\sigma_1, \sigma_2\}$ is $or(X_1, X_2)$-covering.

Stop applies since $X_2$ and 1 are terms of $T(c, N)$.

Finally, considering $and(x_1, x_2)$, we get:

$\text{and}(x_1, x_2) \quad A = T \quad C = T$

Abstract

$\text{and}(X_1, X_2) \quad A = (x_1|_c = X_1 \land x_2|_c = X_2)
C = (\text{and}(x_1, x_2) > x_1, x_2)$

Narrow

$0
\sigma_1 = (X_1 = 0)
A = (x_1|_c = 0 \land x_2|_c = X_2)
C = (\text{and}(X_1, X_2) > x_1, x_2)$

$X_2
\sigma_2 = (X_1 = 1)
A = (x_1|_c = 1 \land x_2|_c = X_2)
C = (\text{and}(X_1, X_2) > x_1, x_2)$

not(not($\text{and}(1, X_2)$))

$\sigma_3 = (X_1 = 1)
A = (x_1|_c = 1 \land x_2|_c = X_2)
C = (\text{and}(X_1, X_2) > x_1, x_2)$

Stop (three times)

$\emptyset
A = (x_1|_c = 0 \land x_2|_c = X_2)
C = (\text{and}(X_1, X_2) > x_1, x_2)$

$\emptyset
A = (x_1|_c = 1 \land x_2|_c = X_2)
C = (\text{and}(X_1, X_2) > x_1, x_2)$

$\emptyset
A = (x_1|_c = 1 \land x_2|_c = X_2)
C = (\text{and}(X_1, X_2) > x_1, x_2)$

$\Sigma = \{\sigma_1, \sigma_2, \sigma_3\}$ is $\text{and}(X_1, X_2)$-covering.

Stop applies on the two first branches because 0 and $X_2$ are terms of $T(c, N)$, and on the third one because it is issued from the same narrowing substitution as the second one, on which Stop applies.

The ordering constraints of the last two proof trees are satisfied by the same ordering as previously.

Note that, as $R$ does not contain rules whose left-hand side is a $C$-term, $C$-reducibility ensures weak termination of the RS.
Now, if we add to $\mathcal{R}$ the relations between constructors

$$
\begin{align*}
1 & \rightarrow \text{true} \\
\text{true} & \rightarrow 1 \\
0 & \rightarrow \text{false} \\
\text{false} & \rightarrow 0
\end{align*}
$$

where $\text{true}$ and $\text{false}$ are also constructor symbols, we loose the weak termination property. But $\mathcal{C}$-reducibility is proved in the same way than previously, since we have no additional proof tree, and the new proof trees are the same as the previous ones.

Let us now consider the second example of the introduction.

**Example:**

$$
\begin{align*}
\text{quot}(0, s(y), s(z)) & \rightarrow 0 \\
\text{quot}(s(x), s(y), z) & \rightarrow \text{quot}(x, y, z) \\
\text{quot}(x, 0, s(z)) & \rightarrow \text{quot}(x, s(z), s(z)) \\
\text{quot}(x, y, 0) & \rightarrow \text{error} \\
\text{quot}(\text{error}, y, z) & \rightarrow \text{error} \\
\text{quot}(x, \text{error}, z) & \rightarrow \text{error} \\
\text{quot}(x, y, \text{error}) & \rightarrow \text{error}
\end{align*}
$$

with $\mathcal{F} = \{ \text{quot} : 3, s : 1, 0 : 0, \text{error} : 0 \}$ and $\mathcal{C} = \{ s, 0, \text{error} \}$.

So, either we prove innermost termination of the RS, (the dependency pair method [18] works), and we infer sufficient completeness, by Proposition 2 and Theorem 1, in proving coveredness of all patterns $f(X_1, \ldots, X_m), f \in \mathcal{D}$, or we directly apply the $\mathcal{C}$-reducibility proof method. Let us develop the second solution.

Applying the rules on $\text{quot}(x_1, x_2, x_3)$, we get:

$$
\begin{align*}
\text{quot}(x_1, x_2, x_3) & \rightarrow A = \top, \quad C = \top \\
\text{Abstract} & \\
\text{quot}(X_1, X_2, X_3) & \rightarrow A = (x_1 | c = X_1 \land x_2 | c = X_2 \\
& \land x_3 | c = X_3) \\
& \land C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3)
\end{align*}
$$

**Narrow**

$$
\begin{align*}
\sigma_1 & = (X_1 = 0 \land X_2 = s(X'_2) \\
& \land X_3 = s(X'_3)) \\
A & = (x_1 | c = 0 \land x_2 | c = s(X'_2) \\
& \land x_3 | c = s(X'_3)) \\
& \land C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3)
\end{align*}
$$

$$
\begin{align*}
\sigma_2 & = (X_1 = s(X'_1) \land X_2 = s(X'_2) \\
A & = (x_1 | c = s(X'_1) \\
& \land x_2 | c = s(X'_2) \land x_3 | c = X_3)) \\
& \land C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3)
\end{align*}
$$

Applying the rule **Narrow** here gives seven branches, following the seven rules of $\mathcal{R}$. Let $u = \text{quot}(X_1, X_2, X_3)$. The set $\Sigma = \{ \sigma_i, i \in [1..7] \}$ is u-covering, since every $\sigma_i$ of $\Sigma$, where $P = \{ 0, \text{error}, s(X)|X \in \mathcal{N} \}$, has a generalization in $\Sigma$.

Then **Stop** applies on all branches, except the third one, for the following reasons. On the second branch and the last four ones, we get $\mathcal{C}$-terms as current terms.

On the second branch, we have $\text{quot}(x_1, x_2, x_3) \succ \text{quot}(X'_1, X'_2, X'_3)$ for a Lexicographic Path Ordering with any precedence and a left-to-right status for $\text{quot}$. Indeed, $\mathcal{R}$ does not contain any rule whose left-hand side is a $\mathcal{C}$-term. Then, if $x_1$ represents a $\mathcal{C}$-term, it is in normal form, and then $x_1 = x_1 | c$. In $A$, we have $x_1 | c = s(X_1)$. Then $x_1 \succ X'_1$. If $x_2$ represents a term containing a defined symbol, then $x_1 \succ X'_1$ (see Section 5). In a similar way, we obtain $x_2 \succ X'_2$.

Now, if $x_3$ represents a term containing a defined symbol, as previously for $x_1$, we have $x_3 \succ X'_3$. If $x_3$ represents a $\mathcal{C}$-term, we get $x_3 = x_3 | c = X_3$, so $x_3 \succ X_3$.

By definition of the LPO, as we have $x_1 \succ X'_1$, we have to verify that $\text{quot}(x_1, x_2, x_3) \succ X'_1, X_3$. This is true since $x_2 \succ X'_2$ and $x_3 \succ X_3$.

Then we get:

$$
\begin{align*}
\text{Stop (six times)} & \\
\emptyset & \rightarrow A = (x_1 | c = 0 \land x_2 | c = s(X'_2) \\
& \land x_3 | c = s(X'_3)) \\
& \land C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3)
\end{align*}
$$
\( \emptyset \)

\( A = (x_1 \downarrow C = s(X'_1)) \)
\( \land x_2 \downarrow C = s(X'_2) \land x_3 \downarrow C = X_3 \)
\( \land \text{quot}(X'_1, X'_2, X_3) \downarrow C = (X_4) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3, x_4, \text{quot}(X'_1, X'_2, X_3)) \)

\( s(\text{quot}(X_1, s(X'_1)), s(X'_3)) \)
\( A = (x_1 \downarrow C = X_1 \land x_2 \downarrow C = 0 \land x_3 \downarrow C = s(X'_3)) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3) \)

\( \emptyset \)

\( A = (x_1 \downarrow C = X_1 \land x_2 \downarrow C = X_2 \land x_3 \downarrow C = 0) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3) \)

\( \emptyset \)

\( A = (x_1 \downarrow C = \text{error} \land x_2 \downarrow C = X_2 \land x_3 \downarrow C = X_3) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3) \)

\( \emptyset \)

\( A = (x_1 \downarrow C = X_1 \land x_2 \downarrow C = X_2 \land x_3 \downarrow C = \text{error}) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3) \)

\( \emptyset \)

\( A = (x_1 \downarrow C = X_1 \land x_2 \downarrow C = X_2 \land x_3 \downarrow C = \text{error}) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3) \)

\( \emptyset \)

\( A = (x_1 \downarrow C = X_1 \land x_2 \downarrow C = X_2 \land x_3 \downarrow C = \text{error}) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3) \)

\( \emptyset \)

\( A = (x_1 \downarrow C = X_1 \land x_2 \downarrow C = X_2 \land x_3 \downarrow C = \text{error}) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3) \)

\( \emptyset \)

\( A = (x_1 \downarrow C = X_1 \land x_2 \downarrow C = X_2 \land x_3 \downarrow C = \text{error}) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3) \)

\( s(\text{quot}(X'_1, X'_2, s(X'_3))) \)
\( \sigma_2 = (X_1 = s(X'_1)) \land x_1 \downarrow C = \text{error} \land x_2 \downarrow C = 0 \land x_3 \downarrow C = s(X'_3)) \)
\( C = (\text{quot}(x_1, x_2, x_3) > x_1, x_2, x_3) \)

\( \Sigma = \{\sigma_1, \sigma_2, \sigma_3\} \) is covering for \( s(\text{quot}(X_1, s(X'_1)), s(X'_3))) \).

Now \textbf{Stop} applies on the first two branches above since \( s(0) \) and \( s(\text{error}) \) are \( C \)-terms.

Finally, \( \text{quot}(x_1, x_2, x_3) > s(\text{quot}(X'_1, X'_2, s(X'_3))) \) since, as \( \text{quot} \in D \) and \( s \in C \), we have \( \text{quot} > x \) \( s \) for the previous LPO. Indeed, in a similar way than previously, if \( x_1 \) is a \( C \)-term, \( x_1 \downarrow C = s(X'_1) > X'_1 \). If \( x_2 \) contains a defined symbol, \( x_1 \downarrow C = s(X'_2) > X'_2 \). We also get \( x_3 \downarrow C = s(X'_3) > X'_3 \).

By definition of the LPO, as we have \( x_1 \uparrow X'_1 \), we have to verify that \( \text{quot}(x_1, x_2, x_3) > X'_3, s(X'_3) \). This is true since \( x_3 \uparrow C = s(X'_3) > X'_3 \).

So \textbf{Stop} applies on the third branch, which ends the proof.

Other examples can be found in [19], as well as proofs of propositions and theorems.

6. Finding a good derivation chain

Our proof process, as it simulates the rewriting mechanism, gives complete information on a rewriting branch leading to a \( C \)-form. It allows us to extract the exact application of rewrite rules that yields a \( C \)-form. To rewrite a term, it is enough to follow the rewriting scheme simulated by abstraction and narrowing in the proof trees.

We now formalize the use of the proof trees to compute a \( C \)-form for any term.

**Definition 12.** Let \( R \) be a RS proved \( C \)-reducible with Theorem 2. The strategy tree \( ST_f \) associated to \( f \in D \) is the proof tree obtained from the initial state \( \{\{f(x_1, \ldots, x_m)\}\}, T, \top\).

**Definition 13.** Let \( R \) be a RS proved \( C \)-reducible with Theorem 2. Let \( ST = \{ST_f\}_f \in D \) be the set of strategy trees of \( R \) and \( f = \{s_1, \ldots, s_n\} \in T(F) \). Rewriting \( s \) with respect to \( ST \) into a \( C \)-form, \( Cform_{ST}(s) \), is defined in the following way, where \( \leftrightarrow \) denotes a transformation step:

- if \( f \in C \) then \( Cform_{ST}(f(s_1, \ldots, s_n)) = f(Cform_{ST}(s_1), \ldots, Cform_{ST}(s_n)) \);
- if \( f \in D \), then reducing \( s \) with respect to \( ST \) into \( Cform_{ST}(s) \) is performed by following the steps in the strategy tree \( ST_f \) of \( f \), where \( t \) is any term of the transformation chain of \( s \) with respect to \( ST \) and \( u \) is the corresponding term in \( ST_f \):
  - if the step is Abstract, and abstracts \( u \) at positions \( s_1, \ldots, s_p \),
    then \( t \mapsto t[\{t'_i\}_{i=1, \ldots, p}] \),
    where \( t'_i = \left\{ \begin{array}{ll}
    t_i & \text{if } C \bis RED(u), \\
    Cform_{ST}(t_i) & \text{otherwise}
    \end{array} \right. \)
  - if the step is Narrow with \( u \mapsto p \downarrow C \downarrow \sigma \ u' \), then \( t \mapsto t' \),
    where \( t' \) is defined by \( t \mapsto p \downarrow C \downarrow \sigma \ t' = \beta t' \), with \( \alpha = \beta \sigma \) on \( \text{Var}(u) \cup \text{Dom}(u) \) and \( t = \alpha u \),
  - if the step is Stop, then \( t \mapsto t' \),
    where \( t' = \left\{ \begin{array}{ll}
    t_i & \text{if } C \bis RED(u), \\
    Cform_{ST}(t) & \text{otherwise}
    \end{array} \right. \).

Given a RS \( R \), the previous definition assumes that if the predicate \( C \bis RED \) has been used to prove \( C \)- reducibility of a particular term \( u \) during the proof that \( R \) is \( C \)-reducing, one is able to find a \( C \)-form for ground instances of \( u \). It is obviously the case if \( u \in T(C, N) \). Under this assumption, the following theorem holds.

**Theorem 3.** Let \( R \) be a RS proved \( C \)-reducing with Theorem 2 and \( ST \) its set of strategy trees. Then for any term \( t \in T(F) \), \( Cform_{ST}(t) \) is a \( C \)-form of \( t \) for \( R \).

7. Conclusion and perspectives

In this paper, we have studied the problem of \( C \)-reducibility of ground terms, i.e., reducibility of terms into a constructor form on at least one of their rewriting derivations. This property is interesting from several points of view. It allows us to warrant a completely computed form for every data of rewrite programs, and directly implies sufficient completeness of specifications. In the light of \( C \)-reducibility, termination of a program becomes less important, provided one is able to reach a \( C \)-form, i.e., to compute a result, which gives a criterion to stop the computations.
Our proof method establishes \( C \)-reducibility, and then sufficient completeness, by explicit induction on the \( C \)-reducibility property itself. It simulates rewriting by alternatively using abstraction of subterms, and narrowing, provided all possible ground instances of the narrowed terms are covered. It works on the ground term algebra using as induction relation a noetherian ordering having the constructor subterm property.

It does not require confluence, nor restrictions like absence of relation between constructors or the constructor preserving property. It does not need any termination property either. When there is no constructor left-hand side of rule, it even provides a proof of weak termination. It is constructive in the sense that a computation branch leading to a constructor form can be deduced from the proof trees.

Thanks to the power of induction, the ordering constraints generated by the proof process are often simple and satisfied by the subterm ordering, or by an usual ordering like the Recursive Path Ordering. If not, they can be delegated to automatic ordering constraint solvers. Verifying unsatisfiability of \( A \) is also automatable, with the given sufficient conditions cited in Section 4.3. So we can have an automatic implementation of the proof procedure establishing \( C \)-reducibility. As the proof technique is close to the rewriting process itself, it seems to be promising to extend it to conditional, typed or constrained rewriting.

References


