Improving basic narrowing techniques[1]

Pierre Réty

Centre de Recherche en Informatique de Nancy
BP 239
54506 Vandoeuvre Les Nancy Cedex
France
E-mail: mcvax!inria!crin!rety

Abstract

In this paper, we propose a new and complete method based on narrowing for solving equations in equational theories. It is a combination of basic narrowing and narrowing with eager reduction, which is not obvious, because their naive combination is not a complete method. We show that it is more efficient than the existing methods in many cases, and for that establish commutation properties on the narrowing. It provides an algorithm that has been implemented as an extension of the REVE software.

1 Introduction

Narrowing is a general method to solve equations in equational theories, that was introduced by Slagle [20], and studied by Fay [2, 1] and Hullot [9, 8]. It needs a convergent set of rewrite rules equivalent to the considered equational theory, and returns a complete set of solutions (also called unifiers), i.e. a basis of the set of all the solutions. But this method has drawbacks: it is inefficient and often does not terminate. Implementations are described in [17, 10]. Narrowing has some similarities with linear resolution principle of the Prolog language, and is used in logic programming language like Eqlog [3] or [21].

Let us describe what narrowing is. Assume that we have a convergent set of rewrite rules. The narrowing of a term t consists of two passes. The first one instantiates t so that it becomes reducible by a rule. The second one reduces it by this rule. The resulting term t_1 may be reducible into t_2 without any further instantiation. If so, t_2 is reduced until one gets a irreducible term (said in normal form) t_n. The relation that transforms t into t_1 was called narrowing in [8, 10], and we call the transformation of t into t_n normalized narrowing. If one considers all the narrowing derivations issued from t (the narrowing tree), the intermediate terms t_1, t_2 are nodes from which edges are issued, while they do not appear in the tree using normalized narrowing. So, the narrowing tree contains the normalized narrowing tree.

In order to compute solutions of an equation $t = t'$ modulo a term rewriting system, one computes the narrowing derivations (normalized or not) issued from $t = t'$, $=$ being considered as a binary function symbol, and check at each node whether the corresponding equation has a syntactic solution. If the term rewriting system is confluent and noetherian, all the solutions are found by building the whole narrowing

[1] This research was supported by the GRECO of programmation.

tree (which can be infinite).
In order to get a smaller tree, it is obviously better to use the normalized narrowing relation. Another idea [8] consists of using only narrowing at some occurrences called basic. This method is called basic narrowing, and gives another tree included into the narrowing tree.

Our idea is to mix the two previous relations, in order to further reduce the tree. The simplest way (we will say naive) is to consider their intersection. Unfortunately, the set of solutions that it provides is not always complete. Therefore, we had to build another relation in a non trivial way that preserves the completeness of the solution set. It use at normalization time a new computation of the basic occurrences based on the residual notion [6], that we will call weakly basic.
In addition this method gives a tree smaller than the one obtained with the normalized narrowing, and smaller than the one obtained with the basic narrowing in many cases. It can be implemented and gives a unification method that is more efficient than the previous ones.

In section 2, we introduce the basic concepts and definitions, and recall the existing results. In section 3, we consider the naive combination of normalized and basic narrowings, and using an example, show that this method does not generate a complete set of solutions. Therefore, we propose in section 4 a new combination of the two relations that provides all the solutions. We compare it with the existing methods in section 6 and for that, we establish commutation results about the narrowing relation in section 5. Details of implementation are described in section 7. The various narrowing relations used in this paper are summarized in the appendix.
All proofs and many examples can be found in [19].

2 Definitions and existing results

In this section we introduce the concept of narrowing and recall that it provides a complete method for solving equation in a theory described by a confluent and noetherian term rewriting system. The following notations and properties are valid for the whole paper. They are consistent with [7, 11].

Let F be a set of symbols, X be a set of variables. A **term** is a partial application from N_+^* (the free monoid on N_+ whose elements are called **occurrences**) into $F \cup X$ that respects the symbol arities. $T(F,X)$ is the set of terms build on F and X. For each $t \in T(F,X)$, $D(t)$ is the set of occurrences of t, $O(t)$ is the set of non variable occurrences and $V(t)$ is the set of variables that occurs in t. t is said **linear** if each variable of t occurs once in t.
An **equation** $s = t$ is a pair of terms, a **rewrite rule** $s \rightarrow t$ is a directed pair of terms satisfying $V(t) \subseteq V(s)$. $t[u \leftarrow t']$ is the term obtained from t by changing the subterm of t at the occurrence u by t'. An **equational theory** A is a set of equations and one writes $=_A$ the smallest congruence induced by A. A **term rewriting system** R is a set of rewrite rules and $\rightarrow$ is the rewriting relation derived from R and $\rightarrow *$ its transitive closure. A sequence of rewriting steps is called a **derivation.** A term t is said normalized if it is not reducible by $\rightarrow$. and the term t' is a normal form of t if $t \rightarrow * t'$ and t' is normalized, t' is also denoted $t\downarrow$. R is **confluent** if for any term t, $t \rightarrow * t_1$ and $t \rightarrow * t_2$ implies there exists a term t' such that $t_1 \rightarrow * t'$ and $t_2 \rightarrow * t'$. R is **noetherian** if the relation $\rightarrow$ is noetherian. R is **interreduced** if for any rule $g \rightarrow d$ in R, d is normalized, and g is normalized with respect to $R - \{g \rightarrow d\}$. One says that R is **convergent** if it is confluent and noetherian, and **canonical** if it is also interreduced. R is **regular** if for all rule $g \rightarrow d$ in R, $V(d) = V(g)$. $=_R$ is the relation defined by $=_R \; = (\rightarrow \cup \leftarrow)^*$ where $\leftarrow$ is the rewriting relation obtained by reversing the rules of R.

Substitutions σ are defined as endomorphisms on $T(F,X)$ that extend mappings from X to $T(F,X)$ with a finite domain $D(\sigma)$. A substitution σ is denoted by $\{(x_1/t_1), \ldots, (x_n/t_n)\}$.
We write $\leqq$ the subsumption quasi-ordering on $T(F,X)$ defined by: $t \leqq t'$ iff $t' = \sigma(t)$ for a substitution σ (called a match from t to t'). Composition of substitutions σ and ρ is denoted by $\sigma . \rho$, then $(\sigma . \rho)(t) = \sigma(\rho(t))$.

Given an equational theory A, two terms t and t' are said to be A-**unifiable** [15, 5] iff there exists a substitution σ such that $\sigma(t) =_A \sigma(t')$. σ is also called an A-**solution** of the equation $t = t'$. Given a subset V of X, we define $\sigma \leqq_A \sigma'[V]$ iff $\sigma' =_A \sigma'' . \sigma[V]$ for some substitution σ'' (the notation $[V]$ means that the formula is valid for any variable in V). If $V = X$, V is omitted. Γ is a **complete set of** A-**unifiers** of t and t' away from W containing the set V of the variables of t and t' iff:

- for all $\sigma \in \Gamma$, $D(\sigma) \subseteq V$ and $I(\sigma) \cap W = \emptyset$ (The goal of this technical restriction is only to avoid conflict between variables)
- for all $\sigma \in \Gamma$, $\sigma(t) =_A \sigma(t')$
- for all unifiers σ', there exists $\sigma \in \Gamma$ such that $\sigma \leqq_A \sigma'[V]$.
  In addition Γ is said to be **minimal** if it satisfies the further condition: for all σ and $\sigma' \in \Gamma$, $\sigma \leqq_A \sigma'$ implies $\sigma = \sigma'$.

An A-unification algorithm is **complete** if it generates a complete set of A-unifiers. Note that this set may be infinite.

We now give a very general definition of narrowing by introducing any fixed mapping $\neg$ such that $\neg \subseteq \rightarrow *$. One will say that a given derivation $s \rightarrow * s'$ is **compatible** with $\neg$ iff $s \neg s'$.

Definition: We say that t is narrowable to t' at the occurrence u, using the rule $g \rightarrow d$ and with the substitution σ iff

- t|u and g are unifiable by the most general unifier σ
- $t_1 = \sigma(t)[u \leftarrow \sigma(d)]$
- $t_1 \neg t'$

We call this relation **narrowing** and denote it $t \mathrel{-\!\overset{\sim}{\rightarrow}}_{[u, g \rightarrow d, \sigma]} t'$. A sequence of narrowing steps is called a **narrowing derivation.**

This definition is generic because by choosing the mapping $\neg$ one obtains different narrowing relations, in particular the two followings:

- If $\neg$ is the identity then $t_1 = t'$. We have the relation called narrowing by Hullot[8], and that we will call **simple narrowing** or S-**narrowing** and we write $t \mathrel{-\!\overset{\wedge}{\rightarrow}}_{[u, g \rightarrow d, \sigma]} t'$. A sequence of S-narrowing steps is called a **S-narrowing derivation.**
- If $\neg$ is the normalization mapping then t' is in normal form. We have the relation called narrowing by Fay [2]. We propose to call it **normalized narrowing** or N-**narrowing** and we denote it by
  $t \mathrel{-\!\overset{\approx}{\rightarrow}}_{[u, g \rightarrow d, \sigma]} t'$. A sequence of N-narrowing steps is called a **N-narrowing derivation.**

With these notations we have:

$$\left[t \mathrel{-\!\overset{\approx}{\rightarrow}}_{[u, g \rightarrow d, \sigma]} t'\right] \Leftrightarrow \left[t \mathrel{-\!\overset{\wedge}{\rightarrow}}_{[u, g \rightarrow d, \sigma]} t_1 \text{ and } t' = t_1 \downarrow\right]$$

$$\rightarrow \subseteq \mathrel{-\!\overset{\wedge}{\rightarrow}} \qquad\qquad \mathrel{-\!\overset{\sim}{\rightarrow}} \subseteq \mathrel{-\!\overset{\wedge}{\rightarrow}} *$$

If σ is a match from g to t/u, the step $t \rightsquigarrow t_1$ is in fact a rewriting step.

In the following we suppose the mapping $\neg$ fixed, so that the narrowing relation $\rightsquigarrow$ is fixed.

The narrowing relation provides a method to compute a complete set of unifiers of two terms modulo a convergent term rewriting system. The method consists in building all the possible narrowing derivations issued from $t_0 = t_0'$ and to collect the corresponding narrowing substitutions, until we obtain equations $t_n = t_n'$ such that t_n and t_n' are unifiable. The unification problem in the equational theory is then reduced to the narrowing together with the standard unification of terms.

In order to iterate the narrowing process on the two terms, $=$ is considered as a new operator of the equational theory, and the process starts with the term $t_0 = t_0'$. It is obvious that if $t_0 = t_0' \rightsquigarrow^* t$ then t is of the form $t_i = t_i'$.

The following result has been proved by Hullot [9] for the S−narrowing, by C. and H. Kirchner [12, 13] and Réty(et al) [18] for any narrowing relation.

Theorem: Let R be a convergent term rewriting system, t_0 and t_0' be two terms. The set of substitutions σ such that

- there exists a narrowing derivation issued from $t_0 = t_0'$ $t_0 = t_0' \rightsquigarrow_{[\sigma_1]} t_1 = t_1' \rightsquigarrow \ldots \rightsquigarrow_{[\sigma_n]} t_n = t_n'$ such that t_n and t_n' are unifiable by the most general unifier β and that $\beta \cdot \sigma_n \ldots \sigma_1$ is normalized on $V(t_0 = t_0')$
- $\sigma = \beta \cdot \sigma_n \ldots \sigma_1$

is a complete set of R−unifiers of t_0 and t_0'.

Basic S−narrowing was defined and studied by Hullot[9]. It consists of forbiding a reduction at an occurrence brought by the substitution in a previous step.

Definition[Hullot]: Given the derivation

$$t_0 \to_{[u_0, g_0 \to d_0, \sigma_0]} t_1 \to \ldots \to_{[u_{n-1}, g_{n-1} \to d_{n-1}, \sigma_{n-1}]} t_n \qquad (1)$$

and $U_0, \ldots, U_n$ sets of non variable occurrences of $t_0, \ldots, t_n$ respectively. One says that the derivation is **based on** U_0 iff for all i

$$u_i \in U_i$$

$$U_{i+1} = \left[U_i - \{v \in U_i \;/\; u_i \leqq v\}\right] \cup \{u_i \,.\, v \;/\; v \in O(d_i)\}$$

We write $U_{i+1} = B(U_i)$, or $U_{i+1} = B(t_{i+1}, (1))$ where (1) specifies which derivation is considered, or more simply $U_{i+1} = B(t_{i+1})$ if it is not ambiguous. One will say U_{i+1} is the base of t_{i+1}. The occurrences that belong to $U_0, \ldots, U_n$ are said basic.

If it is not ambiguous we will say more simply that this derivation is **basic**, or that this is a **basic derivation**. In the same way, we define the **basic S−narrowing derivations**.

Remark: B is monotonic i.e. $U \subseteq U'$ implies $B(U) \subseteq B(U')$, and preserves the closure by prefix i.e U is closed by prefix implies $B(U)$ is closed by prefix.

The basic method consists in building all the S−narrowing derivations issued from $t_0 = t_0'$ and based on $O(t_0 = t_0')$.

Theorem[Hullot]: The previous theorem is still valid when we restrict to S−narrowing derivations based on $U_0 = O(t_0 = t_0')$.

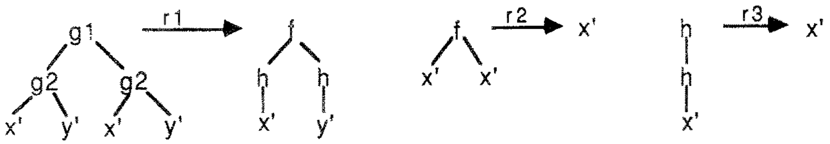
One of the interests of the basic S−narrowing is its termination property.

Termination property[Hullot]: If all the basic S−narrowing derivations issued from a right hand side of a rewrite rule terminate, then all basic S−narrowing derivation issued from any term terminates.

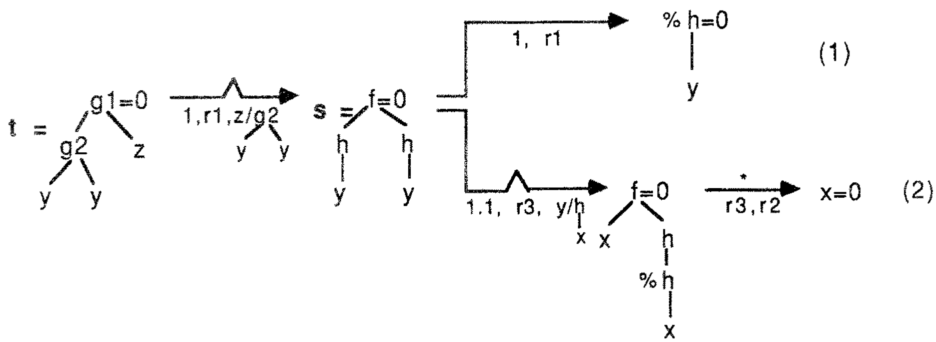
3 The naive basic narrowing

Actually, the "basic" concept is only used for the S−narrowing, and we are loading to extend it to all the narrowing relations (including the N−narrowing). Let us recall that a narrowing step is formed by a step of S−narrowing followed by steps of rewriting. The first idea that comes to mind is to define the basic occurrence sets along the rewriting steps as Hullot did, i.e. by using the mapping B. But this method misses some solutions.

Example 1: Consider the canonical term rewriting system R .



Let $t = g_1(g_2(y,y),z)$ and suppose one wants to solve modulo R the equation $t = 0$. The "%" symbol means that the corresponding occurrence is not basic. $t = 0$ is considered as a term whose top symbol is =. The tree of basic S−narrowing is formed by the branches (1) and (2):



The branch (2) gives the substitution $\sigma = (y / h(0), z / g_2(h(0), h(0)))$ which is the unique solution.

The leaf of the branch (1) can not be narrowed at a basic occurrence, and since $h(y)$ and 0 are not unifiable this branch does not give a solution.

If one uses the naive basic N−narrowing, the term s disappears from the tree, and with it the branch (2). Therefore the solution will not be found by this method.

Nevertheless in order to find σ, our idea is to compute on a larger set of basic occurrences during the rewriting steps. This computing will be said weakly basic. For it the term $h(y) = 0$ can be narrowed into $x = 0$ by the rule r_3 using the substitution $(y / h(x))$ which gives the solution σ.

Another difficulty is that the rewriting steps do not respect the basic occurrences.

Example 2: Let R be the canonical term rewriting system that contains the associativity rule:

$R = \{ f(x, f(y,z)) \to f(f(x,y),z) \}$. Let us apply basic N−narrowing on the term $f(f(y',x'),x')$. One possibility is:

$$f(f(y',x'),x') \leadsto_{[\varepsilon,\sigma]} t := f(f(\%f(y',f(y,z)),y),z)$$
$$\text{with } \sigma = (x'/f(y,z), x/f(y',f(y,z)))$$

The occurrence pointed out by $\%$ is the unique occurrence of t on which the rule can be applied. Since this occurrence is not basic it is not possible to normalize t by a basic derivation.
In the following, we will define a property on basic occurrence sets that will guarantee that basic normalization is possible.

4 The basic narrowing

The aim of the following definition is to characterize the sets of basic occurrences that allow to find a solution. We will say that a set of occurrences U is sufficiently large on a term t if all the subterms that correspond to the non U−occurrences are normalized.

Definition: Let t be a term, U a set of occurrences of t, we say that U is **sufficiently large** on t iff:

$$(u \in D(t) \text{ and } u \notin U) \Rightarrow t/u \text{ is in normal form.}$$

Lemma 1: Let t_0 be a term, U_0 a set of occurrences of t_0 sufficiently large on t_0. Then all the derivations issuing from t_0 and following a bottom-up strategy $t_0 \to t_1 \to \ldots \to t_n$ are based on U_0. If we denote by $U_0, \ldots, U_n$ the sets of basic occurrences then for all $0 \leq i \leq n$, U_i is sufficiently large on t_i.

Proof: By induction on the size of the derivation.
If $n = 0$ the lemma obviously holds. If the property is true for i, U_i is sufficiently large on t_i, then the step $t_i \to_{[u_i, g_i \to d_i, \sigma_i]} t_{i+1}$ satisfies $u_i \in U_i$. Since the strategy is bottom-up, the match σ_i is normalized, and the non basic occurrences of t_{i+1} are normalized. ◆

Corollary 1: If U_0 is sufficiently large on t_0, there exists a derivation based on U_0, leading to the normal form of t_0 and such that for any term t_i in this derivation, the set U_i of the basic occurrences of t_i is sufficiently large on t_i.

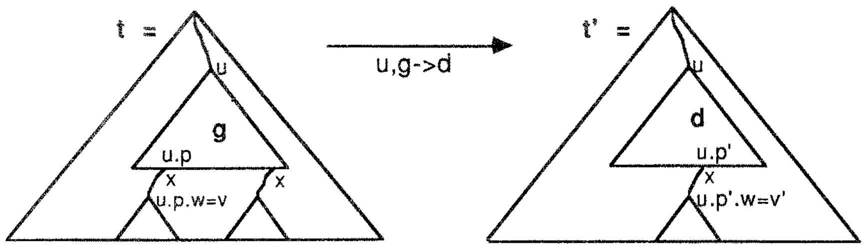
But, there exist basic derivations that do not preserve the sufficient largeness property of the occurrence sets. For instance, consider the rewriting system of the example 1, the term $t = f(h(h(x)), h(h(x)))$, and the occurrence set $U = \{\varepsilon, 1, 11, 2, 21\}$. U is sufficiently large on t, $t \to_{[\varepsilon, r_2]} t' = h(h(x))$ and $B(t') = \emptyset$. Since t' is not normalized, $B(t')$ is not sufficiently large on t'.

We must define a new notion of basic derivation, that always preserves the sufficient largeness property. For that, we introduce the antecedent notion that is (nearly) the dual of the residual notion introduced by Church for the λ−calculus and by Huet and Levy [6] for left−linear term rewriting systems. It characterizes the fact that along a rewriting step, a subterm can be preserved.

Definition: Let $t \to_{[u, g \to d, \sigma]} t'$ be a step of rewriting and $v' \in D(t')$. We say that the occurrence v of t is an **antecedent** of v' iff

$v = v'$ and are not comparable to u
or
there exits an occurrence p' of a variable x in d such that

$v' = u.p'.w$
$v = u.p.w$ where p is an occurrence of x in g.



We extend this definition to a derivation by transitive closure of the rewriting relation. We say that v' is a **residual** of v iff v is an antecedent of v'.

Remarks: With the notations of the previous definition we have:

- $t'/v' = t/v$
- v' may have no antecedent if $v' = u.p'$ with $p' \in O(d)$ or if $v' < u$,
- v' may have several antecedents if g is not linear.

Definition: Given the derivation

$$t_0 \rightarrow_{[u_0, g_0 \rightarrow d_0]} t_1 \rightarrow \cdots \rightarrow_{[u_{n-1}, g_{n-1} \rightarrow d_{n-1}]} t_n$$

and $U_0, \ldots, U_n$ sets of non variable occurrences of $t_0, \ldots, t_n$ respectively. We say that this derivation is **weakly based on** U_0 iff for all i

- $u_i \in U_i$
- $U_{i+1} = \left[U_i - \{v \in U_i \,/\, u_i \leq v\}\right] \cup \{u_i . v \,/\, v \in O(d_i)\}$
  $\cup \{v \in O(t_{i+1}) \,/\, v = u_i . w, w \notin O(d_i)$ and all antecedents of v in t_i are in $U_i\}$

We write $U_{i+1} = WB(U_i)$, $U_{i+1} = WB(t_{i+1}, (1))$ or more simply $U_{i+1} = WB(t_{i+1})$ if it is not ambiguous. One will say U_{i+1} is the base of t_{i+1}. The occurrences that belong to $U_0, \ldots, U_n$ are said basic. If it is not ambiguous we will say more simply that this derivation is **weakly basic**, or that it is a **weakly basic derivation.**

Remark: WB is increasing i.e. $U \subseteq U'$ implies $WB(U) \subseteq WB(U')$, and preserves the closure by prefix i.e U is closed by prefix implies $WB(U)$ is closed by prefix.

This definition differs from Hullot's one by addition of the last line, i.e. the occurrences under d_i may belong to U_{i+1}. Therefore $B(U_i) \subseteq WB(U_i)$.

In practice the set U_0 is supposed to be closed by prefix, and since the weakly basic reduction preserves this property then all the U_i are closed by prefix. Therefore we can get a new and simpler definition of WB by writing:

- $U_{i+1} = \{v \in O(t_{i+1}) \,/\,$ all the antecedents of v in t_i are in $U_i\}$

In the following we will use this definition.

The interest of the weakly basic derivations is pointed out in the following lemma, that emphasizes the fact that the notions of weakly basic derivation and sufficient large occurrence set are very linked.

Lemma 2: Let $t_0 \to * t_n$ be a derivation, and U_0 be a set of occurrences of t_0 sufficiently large on t_0. Then $t_0 \to * t_n$ is weakly based on U_0 and the set U_n of basic occurrences of t_n is sufficiently large on t_n.

Proof: By induction on the length n of the derivation.
If $n = 0$ the lemma obviously holds. Assume $t_0 \to * t_{n-1}$ is weakly basic on U_0 and the basic occurrence set U_{n-1} of t_{n-1} is sufficiently large on t_{n-1}. Thus the reduction occurrence of $t_{n-1} \to t_n$ must be in U_{n-1}, Let U_n be the basic occurrence set of t_n and $v_n \in D(t_n)$ such that $v_n \notin U_n$. From the definition, there exists at least an antecedent v_{n-1} of v_n in t_{n-1} that does not belong to U_{n-1}. Therefore $t_n/v_n = t_{n-1}/v_{n-1}$ which is normalized by hypothesis. ◆

We can now define the basic narrowing with sufficient largeness as a step of basic S−narrowing such that the sufficient largeness property is preserved, followed by a derivation compatible with $\rightharpoondown$. The previous lemma ensures that this derivation is weakly basic, and that the method will be complete.

Definition: Let t_0 be a term, U_0 an occurrence set of t_0, the step of narrowing $t_0 \rightsquigarrow t_n$ (which is equivalent to $t_0 \rightsquigarrow t_1 \rightharpoondown t_n$ i.e. $t_0 \rightsquigarrow t_1 \to \dots \to t_n$) is said **based on U_0 with sufficient largeness** or **SL−based on U_0** iff there are occurrence sets $U_1, \dots, U_n$ such that:

a) $t_0 \rightsquigarrow t_1$ is based on U_0 and $U_1 = B(U_0)$,
b) U_1 is sufficiently large on t_1,
c) For all $i \in \{1, \dots, n-1\}$, $U_{i+1} = WB(U_i)$.

We extend this definition to a narrowing derivation and we will say that a narrowing derivation is SL−based on U_0. If the set U_0 is not specified we will say more simply that this narrowing derivation is SL−basic, or that it is a **SL−basic narrowing derivation.**

This definition prunes the narrowing tree because all the nodes that do not satisfy the sufficient largeness property are cut. As the narrowing definition, this definition is generic and can be instanciated in two particular cases: the **SL−basic S−narrowing** when $\rightharpoondown$ is the identity, and the **SL−basic N−narrowing** when $\rightharpoondown$ the normalization mapping.
The SL−basic S−narrowing and the basic S−narrowing (definition from Hullot) are not the same relations because of the sufficient largeness property imposed by the point b) in the previous definition. The SL−basic S−narrowing relation is included in the basic S−narrowing relation.

Any SL−basic narrowing relation provides a complete method for unifying in a convergent term rewriting system.

Theorem (completeness property): Let R be a convergent term rewriting system, t_0 and t_0' be two terms. The set of substitutions σ such that

- there exists a narrowing derivation issued from $t_0 = t_0'$ and SL−based on $O(t_0 = t_0')$: $t_0 = t_0' \rightsquigarrow_{[\sigma_1]} t_1 = t_1' \rightsquigarrow \dots \rightsquigarrow_{[\sigma_n]} t_n = t_n'$ such that t_n and t_n' are unifiable by the most general unifier β and that $\beta \cdot \sigma_n \dots \sigma_1$ is normalized on $V(t_0 = t_0')$
- $\sigma = \beta \cdot \sigma_n \dots \sigma_1$

is a complete set of R−unifiers of t_0 and t_0'.
Proof: See [19].
Remark: This result still holds if the definition of basic S−narrowing is changed in left−to−right basic S−narrowing defined as in [4], see also [19].

5 Commutation of the narrowing relation

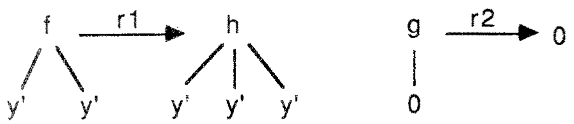
In order to compare the various narrowing relations (next section), we must first study the commutation of the S−narrowing relation with the help of the antecedent notion.
Commutation results are established for the rewriting relation [6] by using the residual notion and for the S−narrowing relation by Herold [4] in the restricted case where the commuted occurrences are not comparable. We establish a more general result by using the antecedent notion.

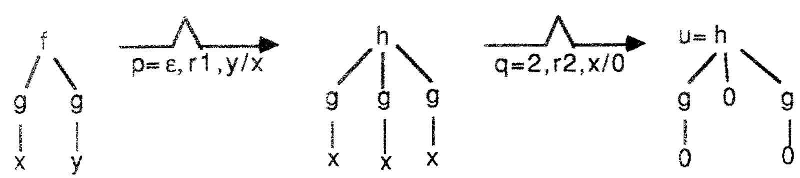
We first extend the antecedent definition on a S−narrowing step.
Definition: Let $t \rightsquigarrow_{[u,g\to d,\sigma]} t'$ be a step of S−narrowing. Let v' be an occurrence of t'. We say that v is an **antecedent** of v' iff

- v is an antecedent of v' in the rewriting step $\sigma(t) \to_{[u,g\to d]} t'$
- $v \in D(t)$.

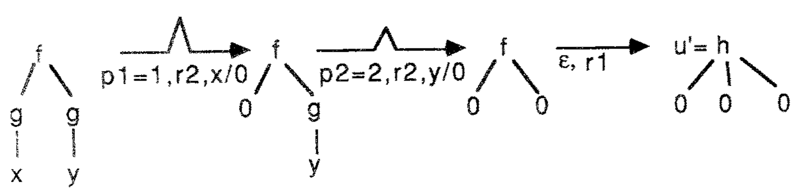
Suppose $s \rightsquigarrow_{[p,g\to d,\sigma]} t \rightsquigarrow_{[q,l\to r,\theta]} u$ (1). One could want to commute the two steps by applying first the rule $l\to r$ and second $g\to d$. If the subterm t/q already existed in s i.e. q admits at least an antecedent in s, the idea consists for applying $l\to r$ in s at all the antecedents of q.

In the following example p is the occurrence of application of the first rule and q those of the second rule.
Example 3: Let the rewriting rules r_1, r_2 be as below and consider the two steps of S−narrowing:





Here $q = 2$ and admits two antecedents $p_1 = 1$ and $p_2 = 2$. Then the two steps commute into three steps:



The leaded terms u and u' are not equal because p_1 and p_2 do not admit only as residual q but also $q_1 = 1$ and $q_2 = 3$. Thus $u \to_{[1,r_2]} \to_{[3,r_2]} u'$.

Others examples and the proof of the following result can be found in [19].
Notations: In order to simplify the notations, we denote the derivation $u \to^*_{[q_1,l\to r]} u_1 \to \ldots \to_{[q_n,l\to r]} u_n = u'$ by $u \to^*_{[q_1,\ldots,q_n,l\to r]} u'$.

Theorem (commutation property): Let R be any term rewriting system and

$$s \mathrel{-\!\hat{\rightarrow}}_{[p, g \to d, \sigma]} t \mathrel{-\!\hat{\rightarrow}}_{[q, l \to r, \theta]} u \qquad (1)$$

be two steps of S−narrowing issued from s such that

- q admits antecedents in s (we denote them by $p_0, \ldots, p_{m-1}$),
- $\theta \cdot \sigma$ is normalized on V(s),
- $V(r) = V(l)$ or g is linear (and in this case m=1).

Then (1) can be commuted into:

$$s \mathrel{-\!\hat{\rightarrow}}_{[p_0, l \to r, \theta'_0]} t'_1 \mathrel{-\!\hat{\rightarrow}} \ldots \mathrel{-\!\hat{\rightarrow}}_{[p_{m-1}, l \to r, \theta'_{m-1}]} t'_m \mathrel{-\!\hat{\rightarrow}}_{[p, g \to d, \sigma']} u' \qquad (2)$$

such that

- $\sigma' \cdot \theta'_{m-1} \ldots \theta'_0 = \theta \cdot \sigma \ [V(s)]$
- $u \to^{*}_{[q_1, \ldots, q_n, l \to r]} u'$ where $q_1, \ldots, q_n$ are the brothers of q i.e. the residuals of $p_0, \ldots, p_{m-1}$ in t.

Remark: If d is linear or u is normalized, or p and q are not comparable, then $u' = u$.

6 Comparison of the narrowing relations

6.1 SL−basic S−narrowing vs basic S−narrowing

Property: SL−basic S−narrowing is strictly included into basic S−narrowing.

Example 8: Consider the canonical term rewriting system

$$R = \{r_1 : f(g(x), y) \to y, r_2 : h(g(x)) \to x, r_3 : f_1(x, x) \to x\}$$

We want to solve modulo R the equation $f_1(0, f(x', h(x'))) = 0$. We compute all the S−narrowing derivations: ("%" symbol means that the corresponding occurrence is not basic)

$$f_1(0, f(x', h(x'))) = 0 \mathrel{-\!\hat{\rightarrow}}_{[1.2, r_1, x'/g(x), y/h(g(x))]} f_1(0, \%h(g(x))) = 0 \qquad (1)$$

$$\mathrel{-\!\hat{\rightarrow}}_{[1.2, r_2, Id]} f_1(0, x) = 0 \mathrel{-\!\hat{\rightarrow}}_{[1, r_3, x/0]} 0 = 0 \qquad (2)$$

and

$$f_1(0, f(x', h(x'))) = 0 \mathrel{-\!\hat{\rightarrow}}_{[1.2.2, r_2, x'/g(x)]} f_1(0, f(g(x), x)) = 0 \qquad (3)$$

$$\mathrel{-\!\hat{\rightarrow}}_{[1.2, r_1, Id]} f_1(0, x) = 0 \mathrel{-\!\hat{\rightarrow}}_{[1, r_3, x/0]} 0 = 0 \qquad (4)$$

The two branches give the solution $x'/g(0)$. One of them is of course useless. By using basic S−narrowing there are only branches (1),(3)+(4); by using SL−basic S−narrowing there are only the branch (3)+(4) because the term $f_1(0, \%h(g(x)))$ contains a non basic subterm (pointed out by "%") that is not in normal form.

6.2 basic narrowing vs basic S−narrowing

In this paragraph, we consider basic narrowing rather than SL−basic narrowing because the sufficient largeness property does not interfere. It is difficult to compare basic narrowing with basic S−narrowing. Indeed, if we consider a basic narrowing derivation, we can transform it into a S−narrowing derivation by considering the rewriting steps as S−narrowing steps. But the rewriting steps use a weakly basic computation of the basic occurrences and then the resulting S−narrowing derivation would not be necessarily basic, but only weakly basic.

Let $t_0 \dashrightarrow^* t_n$ this weakly basic S−narrowing derivation. Suppose the beginning $t_0 \dashrightarrow^* t_i$ is basic, and $t_0 \dashrightarrow^* t_{i+1}$ is not basic. Therefore in the step $t_i \dashrightarrow_{[u_i, g_i \to d_i, \sigma_i]} t_{i+1}$, u_i satisfies $u_i \in WB(t_i)$ and $u_i \notin B(t_i)$. From the definitions there is a step $t_j \dashrightarrow_{[u_j, g_j \to d_j, \sigma_j]} t_{j+1}$ with $j < i$ that creates the difficulty, i.e. the antecedent v_j of u_i in t_j satisfies $v_j \in B(t_j)$ and its antecedent v_{j+1} in t_{j+1} is so that $v_{j+1} \notin B(t_{j+1})$.

In order to transform a weakly basic S−narrowing derivation into a basic S−narrowing derivation, the idea consists of applying the rule $g_i \to d_i$ not on t_i,but on t_j at the occurrence v_j, which belongs to $B(t_j)$. This leads to commute the step $t_i \dashrightarrow t_{i+1}$ with the S−narrowing derivation $t_j \dashrightarrow^* t_i$.

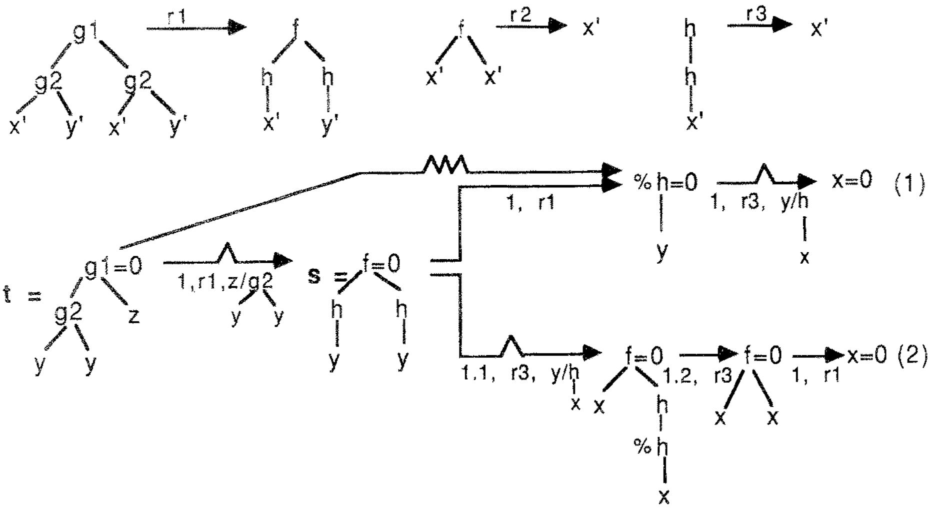
This property is a consequence of the commutation property.

Theorem: Let R be a right−linear term rewriting system, t_0 be a term and U_0 an occurrence set of t_0. We assume moreover R is regular or left linear.

If the narrowing derivation $t_0 \dashrightarrow^*_{[\theta]} t_n$ with θ normalized on $V(t_0)$ is based on U_0, then there exists a S−narrowing derivation $t_0 \dashrightarrow^*_{[\theta]} t_n$ using the same rules and based on U_0.

Therefore, in some cases the basic narrowing is included in the basic S−narrowing. It will then be more interesting to use the basic narrowing, and particularly the basic N−narrowing.

Example 4: Consider the example used for the naive basic narrowing (example 1):



By using basic N−narrowing one obtains the branch (1). One can consider it as a S−narrowing derivation, but then the last step of (1) is not basic, it is only weakly basic. However, by commuting (1) into (2) one obtains an equivalent basic S−narrowing derivation.

6.3 SL−basic N−narrowing vs SL−basic S−narrowing

Under good hypothesis, SL−basic N−narrowing is included in SL−basic S−narrowing, more precisely:

Property: Let R be a right linear term rewriting system. We assume

- R is left linear or for each rule $l \to r$ of R, $V(r) = V(l)$,

- R has no critical pair.

Then if $t_0 \overset{\sim}{-\!\!\rightarrow}{}^*{}_{[\theta]}\, t_n$ by a SL−basic N−narrowing derivation, then there exists a SL−basic S−narrowing derivation $t_0 \overset{\wedge}{-\!\!\rightarrow}{}^*{}_{[\theta']}\, t_n$ such that $\theta' = \theta \ [V(t_0)]$.

7 Implementation

We have proved in the previous section that the SL−basic N−narrowing relation is the smallest narrowing relation provided some conditions are satisfied. We have implemented it within an experimental version of the rewriting software REVE [14] as a modification of the procedure NARROWER [17]. In order to mark the basic occurrences, we have bound a boolean to each occurrence of term, that we call occurrence indicator. Now our implementation does not check the sufficient largeness property and considers that an occurrence is basic if the most left antecedent is basic. Let us describe what will be the final implementation.

Computation of the basic occurrence sets: let us consider a step of weakly basic reduction $t \rightarrow_{[u, g \rightarrow d, \sigma]} t'$ and let us show how the basic occurrences of t' are computed. We have $t' = t[u \leftarrow \sigma(d)]$. Let x be a variable of d that appears at occurrence v'. x appears n times in g at occurrences $v_1, \ldots, v_n$. When the matching process builds the occurrence w of $\sigma(x)$, the occurrences $u.v_1.w, \ldots, u.v_n.w$ of t are examined, and the occurrence indicator of w in $\sigma(x)$ is set to the boolean product of those of $u.v_1.w, \ldots, u.v_n.w$. When g is linear this computation is not more costly than a basic computation.

Test of the sufficient largeness property: consider a basic S−narrowing step $t \overset{\wedge}{-\!\!\rightarrow}_{[u, g \rightarrow d, \sigma]} t'$, In order to check whether the basic occurrence set U' of t' is sufficiently large, before building t' we check that the substitution σ is normalized on $[V(t) \cup V(d)]$. If it is the case, we build $\sigma(t)$ for building t', and verify that $\sigma(t)$ is normalized at all the non basic occurrences of t. Actually, we only test occurrences appearing below some depth, since we know that t is in normal form at the non basic occurrences. Otherwise, the subterms at the non basic occurrences of t' would be normalized, which is further more expensive, therefore this test improves the efficiency.

8 Conclusion

As languages like Eqlog and Slog show, narrowing is a fundamental mechanism for languages that capture both logic and functional programming concepts. The new narrowing relations introduced in this paper allow us to expect more efficient implementations.

Currently, we work in three directions: to make experiments, to study the termination of these new relations, and to extend them to equational rewriting systems.

Appendix: denomination of the various narrowings

A reduction is a sequence of rewriting steps, a normalization is a reduction that leads to the normal form.

narrowing relation	definition	denoted in the literature by:
simple narrowing or S−narrowing (denoted by $-\overset{\wedge}{\rightarrow}$)	more general instantiation and reduction by one rule	narrowing [8]
narrowing (denoted by $-\overset{\sim}{\rightarrow}$)	step of S−narrowing followed by a given reduction	
normal narrowing or N−narrowing (denoted by $-\overset{\approx}{\rightarrow}$)	step of S−narrowing followed by a normalization	narrowing [2, 1], narrowing with eager reduction [16, 10]
weakly basic S−narrowing	S−narrowing with respect to occurrences obtained by a weakly basic computation	
basic S−narrowing	S−narrowing with respect to occurrences obtained by a basic computation	basic narrowing [8]
basic narrowing	step of basic S−narrowing followed by a given and weakly basic reduction	
basic N−narrowing	step of basic S−narrowing followed by a weakly basic normalization	
SL−basic S−narrowing	step of S−narrowing such that the leaded term satisfies the sufficient largeness property	
SL−basic narrowing	step of SL−basic S−narrowing followed by a given and weakly basic reduction	
SL−basic N−narrowing	step of SL−basic S−narrowing followed by a weakly basic normalization	

References

[1] M. Fay: "First−Order Unification in An Equational Theory", *Master Thesis, U. of California At Santa Cruz* , Tech. Report 78−5−002, May 1978.

[2] M. Fay: "First−Order Unification in an Equational Theory", *Proceedings of the 4th Workshop on Automated Deduction* , pp. 161−167, Austin, Texas, 1979.

[3] J. Goguen and J. Meseguer: "EQLOG: Equality, types, and generic Modules for logic programming", *in Functional and Logic Programming* , D. DeGroot and G. Lindstrom, editors, Springer−Verlag, 1985.

[4] A. Herold: "Narrowing techniques applied to idempotent unification", *SEKI report SR−86−16* , August 1986.

[5] G. Huet: "Résolution d'équations dans les langages d'ordre 1,2,...ω", *Thèse d'état* , Université de Paris VII, 1976.

[6] G. Huet and JJ. Levy: "Call by need Computations in Non-ambiguous Linear Term Rewriting Systems", *INRIA Laboria, report 359* , 1979.

[7] G. Huet and D. Oppen: "Equations and Rewrite Rules: A Survey", *in Formal Language Theory: Perspectives and Open Problems* , ed. Book R, Academic Press, 1980.

[8] J.M. Hullot: "Canonical Forms And Unification", *Proceedings of the Fifth Conference on Automated Deduction* , Lecture Notes in Computer Science, vol 87, pp. 318-334, Springer-Verlag, Les Arcs, France, July 1980.

[9] J.M. Hullot: "Compilation de Formes Canoniques dans les Théories Equationnelles", *Thèse de 3ème Cycle* , Université de Paris Sud, 1980.

[10] A. Josephson and N. Dershowitz: "Efficient Implementation of Narrowing: the RITE way", *in Proceedings of the International Conference of Logic Programming* , Salk Lake City, 1986.

[11] J-P. Jouannaud, C. Kirchner, H. Kirchner: "Incremental Construction of Unification Algorithms in Equational Theories", *in Proceedings of the International Conference On Automata, Languages and Programming* , Lecture Notes in Computer Science, vol. 154, pp. 361-373, Springer-Verlag, Barcelona, 1983.

[12] C. Kirchner: "Méthodes et outils de conception systématique d'algorithmes d'unification dans les théories équationnelles", *Thèse de doctorat d'Etat* , Université de Nancy I, 1985.

[13] H. Kirchner: "Preuves par complétion dans les variétés d'algèbres", *Thèse de doctorat d'Etat* , Université de Nancy I, 1985.

[14] P. Lescanne: "Computer experiments with the REVE term rewriting system generator", *in 10th ACM Conf. on Principles of Programming Languages* , pp. 99-108, Austin Texas, January 1983.

[15] G. Plotkin: "Building-in Equational Theories", *in Machine Intelligence* , vol. 7, pp. 73-90, 1972.

[16] U.S. Reddy: "On the relationship between logic and functional languages", *in Logic Programming: Relations, Functions, and Equations* , D. DeGroot and G. Lindstrom, eds. Prentice Hall, Englewood Cliffs, NJ, 1985.

[17] P. Réty, C. Kirchner, H. Kirchner, and P. Lescanne: "NARROWER: A new Algorithm for Unification and its application to Logic Programming", *Proc. 1rst Conf. on Rewriting Techniques and Applications* , Lecture Notes in Computer Science,vol. 202, pp. 141-157, Springer-Verlag, Dijon France, 1985.

[18] P. Réty, C. Kirchner, H. Kirchner, and P. Lescanne: "NARROWER: An algorithm for unification based on narrowing", *CRIN report 86-R-131* , 1986.

[19] P. Réty: "Improving basic narrowing and commutation properties", *CRIN report 87-R-012* , 1987.

[20] J.R. Slagle: "Automated theorem-proving for theories with simplifiers, commutativity and associativity", *J. of ACM* , vol 21, pp. 622-642, 1974.

[21] G. Smolka: "Order-sorted Horn logic semantics and deduction", *Universitat Kaiserslautern* , Internal report, September 1986.