Narrowing Strategies for Arbitrary Canonical Rewrite Systems

Alexander Bockmayr
Stefan Krischer
Andreas Werner

MPI–I–94–233 July 1993

Authors' Addresses

Alexander Bockmayr, Max-Planck-Institut für Informatik, Im Stadtwald, D-66123 Saarbrücken, bockmayr@mpi-sb.mpg.de
Stefan Krischer, CRIN & INRIA-Lorraine, B.P.239, F-54506 Vandœuvre-lès-Nancy, krischer@loria.fr
Andreas Werner, SFB 314, Univ. Karlsruhe, D-76128 Karlsruhe, werner@ira.uka.de

Publication Notes

The present report has been submitted for publication elsewhere and will be copyrighted if accepted.

Acknowledgements

The first author's work was supported by the German Ministry for Research and Technology (BMFT) under grant ITS 9103.
The second author's work was supported by the Centre National de Recherche Scientifique (CNRS), departement Science pour l'Ingenieur.
The third author's work was supported by the Deutsche Forschungsgemeinschaft as part of the SFB 314 (project S2).

Abstract

Narrowing is a universal unification procedure for equational theories defined by a canonical term rewriting system. In its original form it is extremely inefficient. Therefore, many optimizations have been proposed during the last years. In this paper, we present the narrowing strategies for arbitrary canonical systems in a uniform framework and introduce the new narrowing strategy LSE narrowing. LSE narrowing is complete and improves all other strategies which are complete for arbitrary canonical systems. It is optimal in the sense that two different LSE narrowing derivations cannot generate the same narrowing substitution. Moreover, LSE narrowing computes only normalized narrowing substitutions.

Contents

1 Introduction

Narrowing is known as a complete unification procedure for any equational theory that can be defined by a canonical term rewriting system [Fay79, Hul80]. It is also the operational semantics of various logic and functional programming languages. In its original form, narrowing is extremely inefficient [Boc86]. Therefore, many optimizations have been proposed during the last years [Hul80, RKKL85, Fri85, Her86, Rét87, NRS89, BGM88, Ech88, Rét88, Boc88, You88, Pad88, Höl89, DG89, You91].

In this paper, we present the narrowing strategies for arbitrary canonical term rewrite systems in a uniform framework and introduce the new narrowing strategy *LSE narrowing* together with its normalizing variant *normalizing LSE narrowing*. LSE narrowing is complete and improves all previously known strategies which are complete for arbitrary canonical systems, such as left-to-right basic narrowing and the sufficient largeness condition of [Rét87]. It is optimal in the sense that two different LSE narrowing derivations cannot generate the same narrowing substitution. Moreover, there is a one-to-one correspondence between LSE narrowing derivations and a special form of leftmost-innermost rewriting derivations. Finally, LSE narrowing computes only normalized narrowing substitutions.

We are interested in *arbitrary* canonical term rewriting systems that do not have to satisfy additional properties such as constructor discipline [Fri85], left-linearity or non-overlapping left-hand sides [You88, You91, DG89]. For special classes of term rewrite systems, narrowing strategies which are not complete in the general case may be more efficient than LSE narrowing.

For arbitrary canonical systems, the most efficient complete narrowing strategy known before was normalizing SL left-to-right basic narrowing [Rét87]. An analysis of Réty's approach shows that it can be considerably improved if the term rewriting system has non-regular rules and overlapping left-hand sides. In this case various redundancies in the narrowing process can be avoided. LSE narrowing uses three reducibility tests to detect redundant narrowing derivations. The three tests are more powerful than Réty's test for sufficient largeness. Moreover, they imply that any LSE narrowing derivation is also a SL left-to-right basic narrowing derivation. The converse, however, is not true.

The organization of the paper is as follows. After some preliminaries in Section 2, we recall in Section 3 the basic idea of narrowing and give a detailed proof of the well-known lifting lemma of Hullot [Hul80] which establishes a fundamental relationship between rewriting and narrowing derivations. In Section 4, we discuss basic narrowing, left-to-right basic narrowing, and SL left-to-right basic narrowing. While a leftmost-innermost rewriting derivation always generates a SL left-to-right basic narrowing derivation, the converse is not true. In Section 5, we introduce the narrowing strategy *LSE narrowing* and show that there is a one-to-one correspondence between LSE narrowing derivations and left reductions, which are a special form of leftmost-innermost rewriting derivations. Using this correspondence, we can give very simple proofs of the completeness of LSE narrowing and the optimality property that no narrowing substitution can be generated twice. Moreover, we show that LSE narrowing generates only normalized narrowing substitutions. In Section 6, we present the normalizing form of LSE narrowing. The same results hold as in the non-normalizing case. The proofs, however, are more complicated. Finally, in Section 7, we present some empirical results which illustrate the various strategies.

This paper is the full version of [BKW92]. It unifies and simplifies our previous results in [KB91] and [Wer91].

2 Preliminaries

We recall briefly some basic notions that are needed in the sequel. More details can be found in the survey of [HO80].

$\Sigma = (\mathcal{S}, F)$ denotes a *signature* with a set $\mathcal{S}$ of sort symbols and a set F of function symbols together with an arity function.

A Σ-*algebra* A consists of a family of non-empty sets $(A_s)_{s\in\mathcal{S}}$ and a family of functions $(f^A)_{f\in F}$ such that if $f : s_1 \times \ldots \times s_n \to s$ then $f^A : A_{s_1} \times \ldots \times A_{s_n} \to A_s$.

X represents a family $(X_s)_{s\in\mathcal{S}}$ of countably infinite sets X_s of *variables* of sort s. $T(F, X)$ is the Σ-algebra of *terms* with variables over Σ.

For a term $t \in T(F, X)$, $Var(t)$, $Occ(t)$, and $FuOcc(t)$ denote the set of *variables*, *occurrences* and *non-variable occurrences* in t respectively. The root of a term is denoted by the empty occurrence ϵ. An occurence ω is a *prefix* of an occurrence ω', $\omega \preceq \omega'$, iff there exists $\upsilon \in \mathcal{N}^*$ such that $\omega' = \omega.\upsilon$. We denote by t/ω the *subterm* of t at position $\omega \in Occ(t)$ and by $t[\omega \leftarrow s]$ the term obtained from t by *replacing* the subterm t/ω with the term $s \in T(F, X)$.

A *substitution* $\sigma : X \to T(F, X)$ is a family of mappings $\sigma_s : X_s \to T(F, X)_s, s \in \mathcal{S}$, which are different from the identity id only for a finite subset $Dom(\sigma)$ of X. We do not distinguish σ from its canonical extension to $T(F, X)$. $Im(\sigma) \stackrel{\text{def}}{=} \bigcup_{x\in Dom(\sigma)} Var(\sigma(x))$ is the set of *variables introduced* by σ. If σ is a substitution and V is a set of variables then the *restriction* $\sigma|_V$ of σ to V is defined by $\sigma|_V(x) \;=\; \begin{cases} \sigma(x) & \text{if } x \in Dom(\sigma) \cap V \\ x & \text{else} \end{cases}$.

A *syntactic unifier* of two terms s, t is a substitution σ such that $\sigma(s) = \sigma(t)$. A *most general syntactic unifier* of s and t is a unifier σ of s and t with $Dom(\sigma) \cap Im(\sigma) = \emptyset$ such that for any other unifier τ of s and t there exists a substitution λ with $\lambda \circ \sigma = \tau$.

A binary relation $\to = (\to_s)_{s\in\mathcal{S}}$ on a Σ-algebra A is Σ-*compatible* iff $t_1 \to \upsilon_1, \ldots, t_n \to \upsilon_n$ implies $f^A(t_1, \ldots, t_n) \to f^A(\upsilon_1, \ldots, \upsilon_n)$ for all $t_i, \upsilon_i \in A_{s_i}$ and all $f : s_1 \times \ldots \times s_n \to s$ in F. By $\stackrel{*}{\to}$ we denote the reflexive-transitive closure of $\to$. A *congruence* is a Σ-compatible equivalence relation.

An *equation* is an expression of the form $s \doteq t$ where s and t are terms of $T(F, X)$ belonging to the same sort. A *system of equations* G is an expression of the form $s_1 \doteq t_1 \wedge \ldots \wedge s_n \doteq t_n, n \geq 1$ with equations $s_i \doteq t_i, i = 1, \ldots, n$.

Let E be a set of equations. The *equational theory* $\equiv_E$ associated with E is the smallest congruence $\equiv$ on $T(F, X)$ such that $\sigma(l) \equiv \sigma(r)$ for all equations $l \doteq r$ in E and all substitutions σ. Given two substitutions $\sigma, \tau : X \to T(F, X)$ and a set of variables V we write $\sigma = \tau\ [V]$ iff $\sigma(x) = \tau(x)$, for all $x \in V$, and

$\sigma \equiv_E \tau\ [V]$ iff $\sigma(x) \equiv_E \tau(x)$, for all $x \in V$. *E-subsumption* of substitutions is defined by $\sigma \leq_E \tau\ [V]$ iff there is a substitution λ with $\tau(x) \equiv_E \lambda(\sigma(x))$ for all $x \in V$.

A *rewriting rule* π is an expression of the form $l \to r$ with terms $l, r \in T(F, X)$ of the same sort such that $Var(r) \subseteq Var(l)$ and $l \notin X$. The rule is *regular* iff $Var(l) = Var(r)$. The rule is *left-linear* iff no variable occurs twice in l. A *term rewriting system* R is a set of rewriting rules. The *equational theory* $\equiv_R$ generated by R is obtained by considering for every rule $l \to r$ in R the corresponding equation $l \doteq r$.

The *reduction relation* $\to_R$ associated with R is defined as follows: $s \to_R t$, more precisely $s \to_{[v, l \to r, \tau]} t$, iff there is an occurrence $v \in Occ(s)$ and a rule $l \to r$ in R such that there exists a substitution $\tau : X \to T(F, X)$ with $\tau(l) = s/v$ and $t = s[v \leftarrow \tau(r)]$. R is *confluent* iff for any terms s, t_1, t_2 with $s \overset{*}{\to}_R t_1$ and $s \overset{*}{\to}_R t_2$ there exists a term u with $t_1 \overset{*}{\to}_R u$ and $t_2 \overset{*}{\to}_R u$. R is *noetherian* iff there exists no infinite chain $t_1 \to_R t_2 \to_R \ldots \to_R t_n \to_R \ldots$. R is *canonical* iff R is confluent and noetherian.

A term t is *irreducible* or *normalized* iff there exists no term u such that $t \to_R u$. Otherwise t is called *reducible*. A substitution σ is *normalized* iff for any $x \in X$ the term $\sigma(x)$ is irreducible. If R is canonical then there exists for any term t a unique irreducible term $t{\downarrow}$ such that $t \overset{*}{\to}_R t{\downarrow}$. $t{\downarrow}$ is called the *normal form* of t. For any two terms s, t, we have $s \equiv_R t$ iff $s{\downarrow} = t{\downarrow}$.

3 Narrowing: The Basic Idea

Narrowing provides a complete E-unification procedure for any equational theory E that can be defined by a canonical term rewrite system.

Definition 3.1 Let E be a set of equations. A system of equations G

$$s_1 \doteq t_1 \wedge \ldots \wedge s_n \doteq t_n, n \geq 1,$$

is called *E-unifiable* iff there exists a substitution $\sigma : X \to T(F, X)$ such that

$$\sigma(s_1) \equiv_E \sigma(t_1), \ldots, \sigma(s_n) \equiv_E \sigma(t_n).$$

The substitution σ is called an *E-unifier* of G.

A set $cU_E(G)$ of substitutions is called a *complete set of E-unifiers of G* iff

- every $\sigma \in cU_E(G)$ is an E-unifier of G
- for any E-unifier τ of G there is $\sigma \in cU_E(G)$ such that $\sigma \leq_E \tau\ [Var(G)]$
- for all $\sigma \in cU_E(G) : Dom(\sigma) \subseteq Var(G)$.

$cU_E(G)$ is called *minimal* iff it satisfies further the condition

- for all $\sigma, \sigma' \in cU_E(G) : \sigma \leq_E \sigma'\ [Var(G)]$ implies $\sigma = \sigma'$.

Narrowing allows to find complete sets of E-unifiers for equational theories E that can be defined by a canonical term rewrite system R by associating with every rule $l \to r$ in R the equation $l \doteq r$ in E. The basic idea is as follows. Suppose we want to R-unify a system of equations $s_1 \doteq t_1 \wedge \ldots \wedge s_n \doteq t_n$. This means that we have to find a substitution σ such that

$$\sigma(s_1) \equiv_R \sigma(t_1), \ldots, \sigma(s_n) \equiv_R \sigma(t_n). \tag{1}$$

Since R is a canonical term rewriting system this is equivalent to

$$\sigma(s_1){\downarrow} = \sigma(t_1){\downarrow}, \ldots, \sigma(s_n){\downarrow} = \sigma(t_n){\downarrow}. \tag{2}$$

If the problem has a solution σ, then either σ is a syntactic unifier of G, which can be computed by standard unification, or σ does not syntactically unify G. In this case the system of equations $\sigma(G)$ must be reducible by R since otherwise it would be impossible to have (2). The idea is now to *lift* the rewriting derivation $\sigma(G) \;\to \ldots \to\; \sigma(G){\downarrow}$ on the *unknown* instance $\sigma(G)$ of G to a narrowing derivation $G \rightsquigarrow_{\delta_1} \ldots \rightsquigarrow_{\delta_n} G_n$ on the given system G such that the last system of equations G_n is syntactically unifiable with most general unifier τ and $\tau \circ \delta_n \circ \ldots \circ \delta_1 \leq_R \sigma\ [Var(G)]$. This lifting is done by constructing substitutions $\delta_1, \ldots, \delta_n$ such that $\delta_1(G), \ldots, \delta_n(G_{n-1})$ become reducible.

Definition 3.2 (Narrowing) Let R be a term rewriting system. A system of equations G is *narrowable* to a system of equations G' with *narrowing substitution* δ,

$$G \quad \rightsquigarrow_{[v, l \to r, \delta]} \quad G',$$

iff there exist a non-variable occurrence $v \in Occ(G)$ and a rule $l \to r$ in R such that G/v and l are syntactically unifiable with most general unifier δ and $G' = \delta(G)[v \leftarrow \delta(r)]$. We always assume that $Var(l) \cap Var(G) = \emptyset$.

A *narrowing derivation* $G_0 \rightsquigarrow^*_\sigma G_n$ with *narrowing substitution* σ is a sequence of narrowing steps $G_0 \rightsquigarrow_{\delta_1} G_1 \rightsquigarrow_{\delta_2} \ldots \rightsquigarrow_{\delta_n} G_n,\ n \geq 0$, where $\sigma = (\delta_n \circ \ldots \circ \delta_1)\ |_{Var(G)}$. The narrowing substitution leading from G_i to G_j, for $0 \leq i \leq j \leq n$, will be denoted by

$$\lambda_{i,j} \stackrel{\text{def}}{=} \delta_j \circ \ldots \circ \delta_{i+1}.$$

In particular, $\lambda_{i,i} = id$, for $i = 0, \ldots, n$.

A *narrowing strategy* $\mathcal{S}$ is a property of narrowing derivations. We say that *$\mathcal{S}$-narrowing is complete* iff for any canonical term rewriting system R and any system of equations G the set of all substitutions σ such that there exists a $\mathcal{S}$-narrowing derivation $G = G_0 \rightsquigarrow_{\delta_1} \ldots \rightsquigarrow_{\delta_n} G_n, n \geq 0$, such that G_n is syntactically unifiable by a most general unifier τ and $\sigma = \tau \circ \delta_n \circ \ldots \circ \delta_1\ |_{Var(G)}$, is a complete set of R-unifiers of G.

In order to treat syntactical unification as a narrowing step, we introduce a new rule

$$x \doteq x \;\to\; true,$$

where x denotes a variable. Then $t \doteq t' \rightsquigarrow_{\delta} true$ holds if and only if t and t' are syntactically unifiable with most general unifier δ. This additional rule is called *ϵ-rule* (since it can be applied only at occurrence ϵ) and affects neither confluence nor termination. Obviously, σ is a solution of G if and only if $\sigma(G)$ can be reduced by the rules in R and the $\epsilon-$rule to the trivial system $true \wedge \ldots \wedge true$.

Now we are able to formulate the fundamental relationship between rewriting and narrowing derivations that will provide the basis for most of the proofs in this paper.

Proposition 3.3 (Hullot 80) *Let R be a term rewriting system and let G be a system of equations. If μ is a normalized substitution and V a set of variables such that $Var(G) \cup Dom(\mu) \subseteq V$, then for every rewriting derivation*

$$H_0 \stackrel{\text{def}}{=} \mu(G) \quad \rightarrow_{[v_1, l_1 \rightarrow r_1, \tau_1]} \quad H_1 \quad \ldots \quad \rightarrow_{[v_n, l_n \rightarrow r_n, \tau_n]} \quad H_n \qquad (3)$$

there exist a normalized substitution λ and a narrowing derivation

$$G_0 \stackrel{\text{def}}{=} G \quad \rightsquigarrow_{[v_1, l_1 \rightarrow r_1, \delta_1]} \quad G_1 \quad \ldots \quad \rightsquigarrow_{[v_n, l_n \rightarrow r_n, \delta_n]} \quad G_n \qquad (4)$$

using the same rewrite rules at the same occurrences such that

$$\mu = \lambda \circ \delta_n \circ \ldots \circ \delta_1 \quad [V] \quad \textit{and} \qquad (5)$$

$$H_i = (\lambda \circ \delta_n \circ \ldots \circ \delta_{i+1})\,(G_i), \textit{for all } i = 0, \ldots, n. \qquad (6)$$

Conversely, if $\mu \stackrel{\text{def}}{=} \lambda \circ \delta_n \circ \ldots \circ \delta_1$ then there exists for any narrowing derivation (4) *and any substitution λ a rewriting derivation* (3) *such that* (6) *holds.*

Proof: The proof is similar to [Hul80] and uses induction on n. The technical details, however, are different.

First assume that we are given the rewriting derivation (3) of length n.

If $n = 0$, we can choose $\lambda = \mu$.

So suppose $n > 0$ and $\mu(G) \rightarrow_{[v_1, l_1 \rightarrow r_1, \tau_1]} H_1 \rightarrow \ldots \rightarrow H_n$. Then $\mu(G)/v_1 = \tau_1(l_1)$ and $H_1 = \mu(G)[v_1 \leftarrow \tau_1(r_1)]$. Since $\mu(x)$ is irreducible for all $x \in V$ we get $v_1 \in FuOcc(G)$ and $\mu(G)/v_1 = \mu(G/v_1)$. Since we may assume that V and l_1 have no variables in common and that $Dom(\tau_1) \subseteq Var(l_1)$, the substitution $\phi \stackrel{\text{def}}{=} \tau_1 \uplus \mu \stackrel{\text{def}}{=} \begin{cases} \mu(x), & \text{if } x \in Dom(\mu) \\ \tau_1(x), & \text{if } x \in Dom(\tau_1) \end{cases}$ is well-defined and $\phi(G/v_1) = \phi(l_1)$. This means that ϕ is a syntactic unifier of G/v_1 and l_1. Let δ_1 be a most general syntactic unifier of G/v_1 and l_1 with $Dom(\delta_1) \subseteq Var(G/v_1) \cup Var(l_1)$. Then there exists a substitution ρ with $Dom(\rho) \subseteq (((Dom(\tau_1) \cup Dom(\mu)) \setminus Dom(\delta_1)) \cup Im(\delta_1)$ such that $\phi = \rho \circ \delta_1 \; [V \cup Var(l_1)]$. It follows

$$G \rightsquigarrow_{[v_1, l_1 \rightarrow r_1, \delta_1]} G_1$$

with $G_1 \stackrel{\text{def}}{=} \delta_1(G[v_1 \leftarrow r_1])$.

Next we show that the substitution ρ is normalized. Suppose there exists a variable $x \in Dom(\rho)$ such that $\rho(x)$ is reducible. Since μ is normalized we get $x \in (Dom(\tau_1) \setminus Dom(\delta_1)) \cup Im(\delta_1)$. If $x \in Dom(\tau_1) \setminus Dom(\delta_1)$ then $x \in Var(l_1)$ and since $\delta_1(x) = x$, we get $x \in \delta_1(l_1)$. If $x \in Im(\delta_1)$, then it follows from $Dom(\delta_1) \subseteq Var(G/\upsilon_1) \cup Var(l_1)$ that x occurs in $\delta_1(l_1)$ or $\delta_1(G/\upsilon_1)$. But since $\delta_1(l_1) = \delta_1(G/\upsilon_1)$, in both cases x must occur in $\delta_1(G/\upsilon_1)$. So there exists a variable $y \in G$ such that x occurs in $\delta_1(y)$. Then $\rho(x)$ is a subterm of $(\rho \circ \delta_1)(y)$. This implies that $(\rho \circ \delta_1)(y) = \mu(y)$ is reducible in contradiction to the fact that μ is normalized.

Now $\rho(G_1) = \rho(\delta_1(G[\upsilon_1 \leftarrow r_1])) = (\rho \circ \delta_1)(G[\upsilon_1 \leftarrow r_1]) = \phi(G[\upsilon_1 \leftarrow r_1]) = \phi(G)[\upsilon_1 \leftarrow \phi(r_1)] = \mu(G)[\upsilon_1 \leftarrow \tau_1(r_1)] = H_1 \rightarrow_{[\upsilon_2, l_2 \rightarrow r_2, \tau_2]} H_2 \xrightarrow{*} H_n$, with ρ normalized. Let $V' \stackrel{\text{def}}{=} V \cup Im(\delta_1)$. Then by induction hypothesis there exists a substitution λ and a narrowing derivation

$$G_1 \quad \rightsquigarrow_{[\upsilon_2, l_2 \rightarrow r_2, \delta_2]} \quad \dots \quad \rightsquigarrow_{[\upsilon_n, l_n \rightarrow r_n, \delta_n]} \quad G_n$$

such that $\rho = \lambda \circ \delta_n \circ \ldots \circ \delta_2 \ [V']$ and

$$H_i = (\lambda \circ \delta_n \circ \ldots \circ \delta_{i+1})(G_i),$$

for $1 \leq i \leq n$. By the variable disjointness of V and l_1 we get $\mu = \rho \circ \delta_1 \ [V]$. From $V' = V \cup Im(\delta_1)$ and $\rho = \lambda \circ \delta_n \circ \ldots \circ \delta_2 \ [V']$ we conclude $\rho \circ \delta_1 = \lambda \circ \delta_n \circ \ldots \circ \delta_1 \ [V]$. Together this implies

$$\mu = \lambda \circ \delta_n \circ \ldots \circ \delta_1 \ [V]$$

and in particular

$$H_0 = \mu(G_0) = (\lambda \circ \delta_n \circ \ldots \circ \delta_1)(G_0).$$

The reverse direction is again proved by induction.

The case $n = 0$ is trivial. Assume therefore $n > 0$. Let λ be a substitution and

$$G \quad \rightsquigarrow_{[\upsilon_1, l_1 \rightarrow r_1, \delta_1]} \quad G_1 \quad \rightsquigarrow_{[\upsilon_2, l_2 \rightarrow r_2, \delta_2]} \quad \dots \quad \rightsquigarrow_{[\upsilon_n, l_n \rightarrow r_n, \delta_n]} \quad G_n$$

a narrowing derivation. Consider the substitution $\nu \stackrel{\text{def}}{=} \lambda \circ \delta_n \circ \ldots \circ \delta_2$. Then it follows from the induction hypothesis that $\nu(G_1) = H_1 \rightarrow \ldots \rightarrow H_n$ with $H_i = \lambda \circ \delta_n \circ \ldots \circ \delta_{i+1}(G_i)$, for $i = 1, \ldots, n$. From $G \rightsquigarrow_{[\upsilon_1, l_1 \rightarrow r_1, \delta_1]} G_1$ we get $\delta_1(G/\upsilon_1) = \delta_1(l_1)$ and $G_1 = \delta_1(G[\upsilon_1 \leftarrow r_1]) = \delta_1(G)[\upsilon_1 \leftarrow \delta_1(r_1)]$. This means $\delta_1(G) \rightarrow_{[\upsilon_1, l_1 \rightarrow r_1, \delta_1|_{Var(l_1)}]} G_1$. Since $\rightarrow$ is stable under substitutions we obtain $\mu(G) = (\nu \circ \delta_1)(G) \rightarrow \nu(G_1) = H_1 \xrightarrow{*} H_n$. This proves the proposition. □

Proposition 3.4 *A narrowing strategy $\mathcal{S}$ is complete if for any canonical term rewriting system R, any system of equations G, and any normalized R-unifier μ of G there exists a rewriting derivation $\mu(G) \xrightarrow{*} true \wedge \ldots \wedge true$ such that the corresponding narrowing derivation (according to Proposition 3.3) is a $\mathcal{S}$-derivation.*

Proof: Given R and G let $\mathcal{U}$ denote the set of all substitutions σ such that there exists a $\mathcal{S}$-narrowing derivation $G = G_0 \rightsquigarrow_{\delta_1} \ldots \rightsquigarrow_{\delta_n} G_n, n \geq 0$, such that G_n is of the form $true \wedge \ldots \wedge true$ and $\sigma = (\delta_n \circ \ldots \circ \delta_1) \mid_{Var(G)}$.

We will show that $\mathcal{U}$ is a complete set of R-unifiers of G.

By definition, we have $Dom(\sigma) \subseteq Var(G)$. Suppose that $\sigma \in \mathcal{U}$. Then there exists a narrowing derivation $G = G_0 \rightsquigarrow_{\delta_1} \ldots \rightsquigarrow_{\delta_n} G_n, n \geq 0$, such that G_n is of the form $true \wedge \ldots \wedge true$ and $\sigma = (\delta_n \circ \ldots \circ \delta_1) \mid_{Var(G)}$. By Proposition 3.3, there exists a rewriting derivation $\sigma(G) = H_0 \rightarrow_{[l_1 \rightarrow r_1, v_1]} H_1 \ldots \rightarrow_{[v_n, l_n \rightarrow r_n]} H_n$ such that $H_i = (\delta_n \circ \ldots \circ \delta_{i+1})(G_i)$, for $i = 0, \ldots, n$. In particular, we get $H_n = G_n = true \wedge \ldots \wedge true$ which implies that σ is a R-unifier of G.

Now let μ be an arbitrary R-unifier of G and $\mu^{\downarrow}$ its normal form, that is $\mu^{\downarrow}(x) \stackrel{\text{def}}{=} \mu(x)\downarrow$, for all $x \in X$. Then there exists a rewriting derivation $\mu^{\downarrow}(G) = H_0 \stackrel{*}{\rightarrow} H_n = true \wedge \ldots \wedge true$ such that the corresponding narrowing derivation (according to Proposition 3.3) $G = G_0 \rightsquigarrow_{\delta_1} \ldots \rightsquigarrow_{\delta_n} G_n$, is a $\mathcal{S}$-derivation. Furthermore, there is a substitution λ such that $\mu^{\downarrow} = \lambda \circ \delta_n \circ \ldots \circ \delta_1\ [Var(G)]$. By definition, the substitution $\sigma \stackrel{\text{def}}{=} (\delta_n \circ \ldots \circ \delta_1) \mid_{Var(G)}$ belongs to the set of substitutions $\mathcal{U}$ and $\sigma \leq \mu^{\downarrow} \equiv_R \mu\ [Var(G)]$. □

Putting the two preceding propositions together, we get immediately the completeness of naive narrowing.

Theorem 3.5 *Narrowing is complete.*

We close this section with a technical lemma that we will need in the sequel.

Definition 3.6 Two systems of equations G, G' are *identical up to variable renaming* iff there exist substitutions τ, τ' such that $\tau(G) = G'$ and $\tau'(G') = G$.

Lemma 3.7 *Let G_0 and G'_0 be two systems of equations which are identical up to variable renaming. If $G_0 \rightsquigarrow_{[v, l \rightarrow r, \delta]} G_1$ and $G'_0 \rightsquigarrow_{[v, l \rightarrow r, \delta']} G'_1$, then G_1 and G'_1 are also identical up to variable renaming.*

Proof: Suppose that $\tau_0(G_0) = G'_0$ and $\tau'_0(G'_0) = G_0$, for some substitutions τ_0, τ'_0 with $Dom(\tau_0) \subseteq Var(G_0)$ and $Dom(\tau'_0) \subseteq Var(G'_0)$.

By definition, $G_1 = \delta(G_0[v \leftarrow r])$ and $G'_1 = \delta'(G'_0[v \leftarrow r])$ with a most general unifier δ of G_0/v and l and a most general unifier δ' of G'_0/v and l. Since $G_0/v = \tau'_0(G'_0)/v$ and $G'_0/v = \tau_0(G_0)/v$ we can conclude from $Dom(\tau_0) \cap Var(l) = Dom(\tau'_0) \cap Var(l) = \emptyset$ that $\delta(\tau'_0(l)) = \delta(l) = \delta(G_0/v) = \delta(\tau'_0(G'_0/v))$ and similarly $\delta'(\tau_0(l)) = \delta'(l) = \delta'(G'_0/v) = \delta'(\tau_0(G_0/v))$. Hence $\delta \circ \tau'_0$ unifies G'_0/v and l and $\delta' \circ \tau_0$ unifies G_0/v and l. Since δ and δ' are most general unifiers, there exist substitutions τ_1 and τ'_1 with $\tau_1 \circ \delta = \delta' \circ \tau_0$ and $\tau'_1 \circ \delta' = \delta \circ \tau'_0$. It follows $\tau_1(G_1) = \tau_1(\delta(G_0[v \leftarrow r])) = \delta'(\tau_0(G_0[v \leftarrow r])) = \delta'(\tau_0(G_0)[v \leftarrow \tau_0(r)]) = \delta'(G'_0[v \leftarrow r]) = G'_1$ and similarly $\tau'_1(G'_1) = G_1$. This shows that G_1 and G'_1 are identical up to variable renaming. □

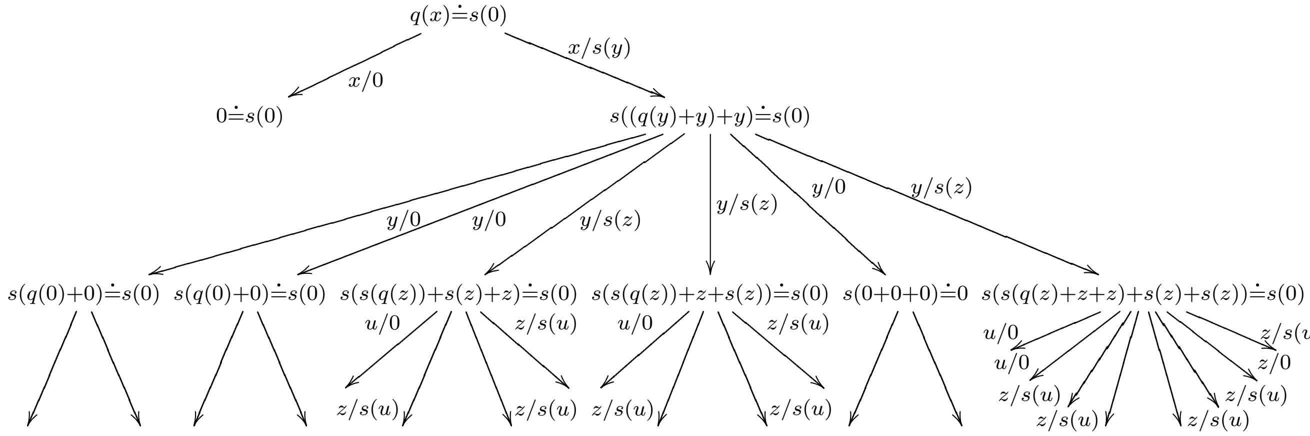

Figure 1: Narrowing tree for $q(x) \doteq s(0)$

4 Left-to-Right Basic Narrowing

Narrowing in its original form is extremely inefficient [Boc86].

Example 4.1 Consider the canonical rewrite system

$$\begin{array}{rlll} R & = \{ & x+0 \rightarrow x, & x+s(y) \rightarrow s(x+y), \\ & & q(0) \rightarrow 0, & q(s(x)) \rightarrow s((q(x)+x)+x)\} \end{array}$$

for adding and squaring natural numbers and suppose we want to solve the query $q(x) \doteq s(0)$. The corresponding narrowing tree is given in Fig. 1.

The example shows that on different pathes in the search tree the same narrowing substitution is generated again and again. The narrowing substitution, however, is the only interesting information obtained in a narrowing derivation. If there are two derivations $G \rightsquigarrow^*_{\sigma_1} G_1$ and $G \rightsquigarrow^*_{\sigma_2} G_2$ such that $\sigma_1 = \sigma_2\ [Var(G)]$, then $G_1 \equiv_R G_2$ independently of the narrowing occurrences and the rules that have been selected. In this case, one of the two derivations is redundant.

If μ is a normalized R-unifier of the system of equations G, then *any* rewriting derivation $\mu(G) \xrightarrow{*} true \wedge \ldots \wedge true$ can be lifted to a narrowing derivation $G \rightsquigarrow^*_{\sigma} true \wedge \ldots \wedge true$ such that the narrowing substitution σ subsumes μ. It follows that there are as many narrowing derivations generating (a generalization of) the same solution μ as there are different normalizations of $\mu(G)$. This is one of the main reasons for the inefficiency of narrowing.

The natural solution to this problem is

- to introduce a normalization strategy for $\mu(G)$ and
- to consider only those narrowing derivations which correspond to rewriting derivations following this strategy.

4.1 Basic Narrowing

A first step in this direction was Hullot's *basic narrowing* [Hul80]. A *basic narrowing derivation* is obtained, when an *innermost* rewriting derivation on $\mu(G)$, where μ denotes a normalized substitution, is lifted to the narrowing level.

Definition 4.2 (Basic Narrowing) The sets $B_i, i = 0, \ldots, n$, of *basic occurrences* in a narrowing derivation

$$G_0 \rightsquigarrow_{[v_1, l_1 \to r_1, \delta_1]} G_1 \rightsquigarrow_{[v_2, l_2 \to r_2, \delta_2]} \cdots \rightsquigarrow_{[v_n, l_n \to r_n, \delta_n]} G_n$$

are inductively defined as follows

- $B_0 \stackrel{\text{def}}{=} FuOcc(G_0)$
- $B_i \stackrel{\text{def}}{=} (B_{i-1} \setminus \{v \in B_{i-1} \mid v \succeq v_i\}) \cup \{v_i.v \mid v \in FuOcc(r_i)\}, i > 0.$

For a *basic narrowing derivation* we require that $v_i \in B_{i-1}$, for all $i = 1, \ldots, n$.

While original narrowing considers any non-variable occurrence in the goal, basic narrowing discards those occurrences which have been introduced by the narrowing substitution of a previous narrowing step.

Since, in canonical systems, an innermost normalization of $\mu(G)$ always exists, Proposition 3.4 implies that basic narrowing is complete. A formal proof will be given in Corollary 5.13. Note that naive innermost narrowing is not complete.

4.2 Left-to-Right Basic Narrowing

In 1986, Herold showed that it is possible to restrict the set of narrowing occurrences further without loosing completeness. After a narrowing step $G \rightsquigarrow_{[v, l \to r, \sigma]} G'$, we may discard also those narrowing occurrences which are strictly left of v [Her86].

Definition 4.3 An occurrence ω is *strictly left* of an occurrence ω', $\omega \lhd \omega'$ (resp. $\omega' \rhd \omega$) iff there exist occurrences o, v, v' and natural numbers i, i' such that $i < i', \omega = o.i.v$ and $\omega' = o.i'.v'$.

Definition 4.4 (Left-to-Right Basic Narrowing) The sets $LRB_i, i = 0, \ldots, n$, of *left-to-right basic occurrences* in a narrowing derivation

$$G_0 \rightsquigarrow_{[v_1, l_1 \to r_1, \delta_1]} G_1 \rightsquigarrow_{[v_2, l_2 \to r_2, \delta_2]} \cdots \rightsquigarrow_{[v_n, l_n \to r_n, \delta_n]} G_n$$

are inductively defined as follows

- $LRB_0 \stackrel{\text{def}}{=} FuOcc(G_0)$
- $LRB_i \stackrel{\text{def}}{=} (LRB_{i-1} \setminus \{v \in LRB_{i-1} \mid v \succeq v_i \text{ or } v \lhd v_i\}) \cup \{v_i.v \mid v \in FuOcc(r_i)\}, i > 0.$

For a *left-to-right basic narrowing derivation* we require that $v_i \in LRB_{i-1}$, for all $i = 1, \ldots, n$. Sometimes, we will use the abbreviation $LRB(U, v, l \to r)$ for the set of occurrences $(U \setminus \{\omega \in U \mid \omega \succeq v \text{ or } \omega \lhd v\}) \cup \{v.\omega \mid \omega \in FuOcc(r)\}$.

We could also define a right-to-left basic narrowing derivation. If we allow arbitrary selection strategies we obtain the *basic selection narrowing* of [BGM88], which includes left-to-right and right-to-left basic narrowing as special cases.

Herold showed that narrowing derivations corresponding to leftmost-innermost normalizations of $\mu(G)$, for a normalized substitution μ, are left-to-right basic. This implies immediately the completeness of left-to-right basic narrowing (see Corollary 5.14 for a formal proof).

4.3 SL Left-to-Right Basic Narrowing

To further improve left-to-right-basic narrowing, Réty introduced the notion of sufficient largeness [Rét87, Rét88].

Definition 4.5 (Sufficient Largeness) A set U of occurrences of a term t is said to be *sufficiently large* on t, iff t/ω is in normal form for all $\omega \in Occ(t) \setminus U$.

Réty noticed that sufficient largeness is preserved by leftmost-innermost rewriting derivations.

Lemma 4.6 *Let $H_0 \xrightarrow{*} H_n$ be a leftmost-innermost rewriting derivation, U_0 a sufficiently large set of occurrences of H_0, and $U_{i+1} \stackrel{\text{def}}{=} LRB(U_i, v_{i+1}, \pi_{i+1})$. Then U_i is sufficiently large on H_i, for all $i = 0, \ldots, n$.*

Proof: By induction on the length of the derivation. For $n = 0$ the lemma is trivial. If U_n is sufficiently large on H_n, then the step $H_n \to_{[v,\pi,\tau]} H_{n+1}$ satisfies $v \in U_n$. Since the strategy is innermost, the matching substitution τ is normalized. This holds because l cannot be a variable and therefore $\tau(x)$ is a proper subterm of $\tau(l)$. Since the strategy is leftmost, the part of H_n strictly left of v is normalized. This shows that U_{n+1} is sufficiently large on H_{n+1}. □

Lifting this property to the narrowing level yields SL left-to-right basic narrowing.

Definition 4.7 (SL Left-to-Right Basic Narrowing) A *SL left-to-right basic narrowing derivation* is a left-to-right basic narrowing derivation

$$G_0 \quad \rightsquigarrow_{[v_1, l_1 \to r_1, \delta_1]} \quad G_1 \quad \rightsquigarrow_{[v_2, l_2 \to r_2, \delta_2]} \quad \cdots \quad \rightsquigarrow_{[v_n, l_n \to r_n, \delta_n]} \quad G_n$$

such that the set of left-to-right basic occurrences LRB_i is sufficiently large on G_i, for $i = 1, \ldots, n$.

By Proposition 3.4, we can conclude that SL left-to-right basic narrowing is complete (see also Corollary 5.15).

While lifting a leftmost-innermost rewriting derivation to the narrowing level always yields a SL left-to-right basic narrowing derivation, the converse is not true.

A SL left-to-right basic narrowing derivation need not generate a leftmost-innermost rewriting derivation.

Example 4.8 Consider the rule

$$\pi : \quad z * 0 \quad \rightarrow \quad 0.$$

Starting with the term $(y * x) * x$ there are two SL left-to-right basic narrowing derivations

$$\begin{array}{llllll} (y * x) * x & \rightsquigarrow_{[1,\pi,\{x \leftarrow 0, z \leftarrow y\}]} & 0 * 0 & \rightsquigarrow_{[\epsilon,\pi]} & 0 \\ (y * x) * x & \rightsquigarrow_{[\epsilon,\pi,\{x \leftarrow 0, z \leftarrow y*0\}]} & & & 0 \end{array}$$

There is an obvious redundancy. In both derivations, the narrowing substitution $\{x \leftarrow 0\}$ and the derived term 0 are the same.

The reduction

$$(y * 0) * 0 \quad \rightarrow_{[\epsilon,\pi]} \quad 0$$

corresponding to the second narrowing derivation is not leftmost-innermost, since $y * 0$ can be reduced.

5 LSE Narrowing

Our aim is now to introduce a new narrowing strategy which has the property that the corresponding rewriting derivations are always leftmost-innermost.

We start by refining the notion of a leftmost-innermost rewriting derivation. Leftmost-innermost derivations are not unique. If the rewrite system has unifiable left-hand sides, then it may happen that two different rules are applicable at the same occurrence. In order to eliminate this indeterminism we assume that the rules are ordered by a total well-founded ordering $<$. If several rules can be applied at the same occurrence, we require that the minimal rule is chosen.

Definition 5.1 (Left Reduction) A reduction step $t \rightarrow_{[\upsilon,\pi,\sigma]} t'$ is called a *left reduction step* iff

- all subterms t/ω with ω strictly left of υ are in normal form ("leftmost")
- all proper subterms of t/υ are in normal form ("innermost")
- t/υ cannot be reduced by a rule π' smaller than π ("minimal rule").

A rewriting derivation is a *left reduction* iff all steps are left reduction steps.

While leftmost-innermost derivations are not unique due to the indeterminism in the selection of the rule, left reductions are unique.

Proposition 5.2 *For all terms t there exists a unique left reduction to the normal form of t.*

Proof: We prove the proposition by noetherian induction. If t is in normal form, then the theorem holds trivially. If t can be reduced, then there exists a unique first left reduction step $t \rightarrow t'$, since the ordering $\lhd \cup \prec$ on $Occ(t)$ and the ordering on rules are total and well-founded. By induction hypothesis, there is a unique left derivation $t' \stackrel{*}{\rightarrow} t'\downarrow$. If we join the two derivations together, we get the unique left reduction $t \rightarrow t' \stackrel{*}{\rightarrow} t'\downarrow = t\downarrow$. $\square$

Now we will show how *reducibility tests* which are performed after a narrowing step can be used to obtain a one-to-one correspondence between narrowing derivations and left reductions.

Definition 5.3 (LSE Narrowing) In a narrowing derivation

$$G_0 \quad \rightsquigarrow_{[\upsilon_1,\pi_1,\delta_1]} \quad G_1 \quad \rightsquigarrow_{[\upsilon_2,\pi_2,\delta_2]} \cdots \quad G_{n-1} \quad \rightsquigarrow_{[\upsilon_n,\pi_n,\delta_n]} \quad G_n$$

the step $G_{n-1} \rightsquigarrow_{[\upsilon_n,\pi_n,\delta_n]} G_n$ is called *LSE* iff the following three conditions are satisfied:

(Left-Test) For all $i \in \{0,\ldots,n-1\}$ the subterms of $\lambda_{i,n}(G_i)$ which lie strictly left of υ_{i+1} are in normal form.

(Sub-Test) For all $i \in \{0,\ldots,n-1\}$ the proper subterms of $\lambda_{i,n}(G_i/\upsilon_{i+1})$ are in normal form.

(Epsilon-Test) For all $i \in \{0,\ldots,n-1\}$ the term $\lambda_{i,n}(G_i/\upsilon_{i+1})$ is not reducible at occurrence ϵ with a rule smaller than π_{i+1}.

A narrowing derivation is *LSE* iff any single narrowing step is LSE.

In [KB91], LSE narrowing was introduced as a refinement of SL left-to-right basic narrowing [Rét87]. A *LSE-SL left-to-right basic narrowing derivation* was defined as a left-to-right basic narrowing derivation for which the SL-Test, the Sub-Test, and the Epsilon-Test detect no redundancy. By introducing the Left-Test, this definition and the subsequent proofs could be considerably simplified. The Left-Test replaces the SL-Test and implies together with the Sub-Test that a LSE narrowing derivation is also left-to-right basic. However, while the notion of left-to-right basic occurrences is not needed anymore in the definition of LSE narrowing, it is still very useful in a practical implementation. We do not have to perform the Left- or Sub-Test at a non-left-to-right-basic occurrence because we know in advance that a redundancy will be detected.

Proposition 5.4 *Consider a system of equations G and a normalized substitution μ. If*

$$H_0 \stackrel{\text{def}}{=} \mu(G_0) \quad \rightarrow_{[\upsilon_1,\pi_1]} \quad H_1 \quad \rightarrow_{[\upsilon_2,\pi_2]} \cdots \rightarrow_{[\upsilon_n,\pi_n]} \quad H_n$$

is a left reduction, then the corresponding narrowing derivation

$$G_0 \quad \rightsquigarrow_{[\upsilon_1,\pi_1,\delta_1]} \quad G_1 \quad \rightsquigarrow_{[\upsilon_2,\pi_2,\delta_2]} \quad \cdots \quad G_{n-1} \quad \rightsquigarrow_{[\upsilon_n,\pi_n,\delta_n]} \quad G_n$$

is a LSE narrowing derivation.

Proof: By Proposition 3.3 there exists a substitution λ such that $H_i = (\lambda \circ \lambda_{i,n})(G_i)$, for $i = 0, \ldots, n$. We have to show that none of the reducibility tests detects a redundancy.

Suppose that the step $G_{m-1} \rightsquigarrow_{[\upsilon_m,\pi_m,\delta_m]} G_m$ is not LSE, for some $m \in \{1, \ldots, n\}$. Then there exists $i \in \{0, \ldots, m-1\}$ such that either

1. $\lambda_{i,m}(G_i)$ is reducible at an occurrence υ strictly left of υ_{i+1} or
2. $\lambda_{i,m}(G_i)$ is reducible at an occurrence υ strictly below υ_{i+1} or
3. $\lambda_{i,m}(G_i)$ is reducible at occurrence υ_{i+1} with a rule smaller than π_{i+1}.

Since $H_i = (\lambda \circ \lambda_{m,n} \circ \lambda_{i,m})(G_i)$ and $\rightarrow$ is stable under substitutions this implies that one of the properties (1) to (3) must hold with H_i in place of $\lambda_{i,m}(G_i)$. But this means that $H_i \rightarrow_{[\upsilon_{i+1},\pi_{i+1}]} H_{i+1}$ is not a left reduction step in contradiction to our assumption. □

As an immediate consequence, we get by Proposition 5.2 and Proposition 3.4 the following theorem.

Theorem 5.5 *LSE narrowing is complete.*

Next we consider the converse of Proposition 5.4.

Proposition 5.6 *If*

$$G_0 \rightsquigarrow_{[\upsilon_1,\pi_1,\delta_1]} G_1 \rightsquigarrow_{[\upsilon_2,\pi_2,\delta_2]} \cdots G_{n-1} \rightsquigarrow_{[\upsilon_n,\pi_n,\delta_n]} G_n$$

is a LSE narrowing derivation and $H_i \stackrel{\text{def}}{=} \lambda_{i,n}(G_i)$, for $i = 0, \ldots, n$, then the rewriting derivation

$$H_0 \rightarrow_{[\upsilon_1,\pi_1]} H_1 \rightarrow_{[\upsilon_2,\pi_2]} \cdots \rightarrow_{[\upsilon_n,\pi_n]} H_n$$

is a left reduction.

Proof: Suppose that the derivation is not a left reduction. Then there exists $i \in \{0, \ldots, n-1\}$ and a rewriting step $\lambda_{i,n}(G_i) \rightarrow_{[\upsilon,\pi]} \lambda_{i+1,n}(G_{i+1})$ such that either

1. υ lies strictly left of υ_{i+1} or
2. υ lies strictly below υ_{i+1} or
3. the rule π is smaller than π_{i+1}

But this implies that the narrowing step $G_{n-1} \rightsquigarrow_{[\upsilon_n,\pi_n,\delta_n]} G_n$ is not LSE in contradiction to our assumption. □

This proposition has a number of important consequences. First of all, we can easily prove the following minimality property of LSE narrowing which first appeared in [Wer91].

Theorem 5.7 *Consider two LSE narrowing derivations*

$$G = G_0 \;\rightsquigarrow_{[\upsilon_1,\pi_1,\delta_1]}\; G_1 \;\rightsquigarrow_{[\upsilon_2,\pi_2,\delta_2]}\; \ldots \; G_{n-1} \;\rightsquigarrow_{[\upsilon_n,\pi_n,\delta_n]}\; G_n,$$

$$G = G'_0 \;\rightsquigarrow_{[\upsilon'_1,\pi'_1,\delta'_1]}\; G'_1 \;\rightsquigarrow_{[\upsilon'_2,\pi'_2,\delta'_2]}\; \ldots \; G'_{m-1} \;\rightsquigarrow_{[\upsilon'_m,\pi'_m,\delta'_m]}\; G'_m.$$

If the narrowing substitutions $\sigma \stackrel{\text{def}}{=} \delta_n \circ \ldots \circ \delta_1$ *and* $\sigma' \stackrel{\text{def}}{=} \delta'_m \circ \ldots \circ \delta'_1$*, where* $n \leq m$*, coincide on* $Var(G)$ *up to variable renaming, that is if there exist substitutions* λ *and* λ' *such that* $\sigma = \lambda' \circ \sigma' \,[Var(G)]$ *and* $\sigma' = \lambda \circ \sigma \,[Var(G)]$*, then*

- $\pi_i = \pi'_i$ *and* $\upsilon_i = \upsilon'_i$ *for* $1 \leq i \leq n$*,*
- *the narrowing derivation*

  $$G'_n \rightsquigarrow_{[\upsilon'_{n+1},\pi'_{n+1},\delta'_{n+1}]} G'_{n+1} \ldots \rightsquigarrow_{[\upsilon'_m,\pi'_m,\delta'_m]} G'_m$$

  is a left reduction (up to variable renaming).

Proof: By Proposition 5.6 the rewriting derivations

$$\sigma(G) = \lambda_{0,n}(G_0) \;\rightarrow_{[\upsilon_1,\pi_1]}\; \lambda_{1,n}(G_1) \;\rightarrow_{[\upsilon_2,\pi_2]}\; \ldots \;\rightarrow_{[\upsilon_n,\pi_n]}\; \lambda_{n,n}(G_n)$$

$$\sigma'(G) = \lambda'_{0,m}(G'_0) \;\rightarrow_{[\upsilon'_1,\pi'_1]}\; \lambda'_{1,m}(G'_1) \;\rightarrow_{[\upsilon'_2,\pi'_2]}\; \ldots \;\rightarrow_{[\upsilon'_m,\pi'_m]}\; \lambda'_{m,m}(G'_m)$$

are both left reductions. Since σ and σ' coincide on $Var(G)$ up to variable renaming, the systems $\sigma(G)$ and $\sigma'(G)$ are identical up to variable renaming. By the unicity of left reductions, this implies $\pi_i = \pi'_i$ and $\upsilon_i = \upsilon'_i$ for $1 \leq i \leq n$.

Using Lemma 3.7 we can conclude by induction that G_i and G'_i resp. $\lambda_{i,n}(G_i)$ and $\lambda'_{i,m}(G'_i)$ are identical up to variable renaming, for $i = 1, \ldots, n$. Since $\lambda_{n,n}(G_n) = G_n$, this implies that G'_n and $\lambda'_{n,m}(G'_n)$ are identical up to variable renaming. Therefore, again by Lemma 3.7, the narrowing derivation starting from G'_n and the left reduction starting from $\lambda'_{n,m}(G'_n)$ are the same up to variable renaming. □

If we assume that narrowing derivations starting from the same goal and using the same rules at the same occurrences produce the same narrowing substitution (in any practical implementation, this will be the case), we get the following corollary.

Corollary 5.8 *If LSE narrowing enumerates two solutions* σ *and* σ' *which coincide up to variable renaming, then* $\sigma = \sigma'$ *holds and the two derivations coincide.*

Proof: We use the same notation as in Theorem 5.7. Then $G_n = G'_m = true \wedge \ldots \wedge true$ implies $n = m$. □

Another important property of LSE narrowing is that it generates only *normalized substitutions*. The other narrowing strategies produce also non-normalized substitutions, which blow up the narrowing search space. If one wants to eliminate them one has to perform an additional normalization test, which is not necessary for LSE narrowing.

Proposition 5.9 *For any LSE narrowing derivation*

$$G = G_0 \quad \rightsquigarrow_{[\upsilon_1,\pi_1,\delta_1]} \quad G_1 \quad \rightsquigarrow_{[\upsilon_2,\pi_2,\delta_2]} \quad G_2 \quad \ldots \quad \rightsquigarrow_{[\upsilon_n,\pi_n,\delta_n]} \quad G_n$$

the narrowing substitution $\delta_n \circ \ldots \circ \delta_1|_{Var(G)}$ *is normalized.*

Proof: Let x be a variable of G such that $\lambda_{0,n}(x)$ is reducible. Suppose x is instantiated for the first time in the i-th narrowing step. Then there must be an occurrence of x in G_{i-1} which lies below the narrowing occurrence υ_i. More formally, there exists an occurrence $\upsilon \neq \epsilon$ such that $G_{i-1}/\upsilon_i.\upsilon = x$. Then $\lambda_{i-1,n}(G_{i-1}/\upsilon_i.\upsilon) = \lambda_{i-1,n}(x) = \lambda_{0,n}(x)$ is reducible in contradiction to the Sub-Test. □

Corollary 5.10 *LSE narrowing enumerates only normalized substitutions.*

Note that the last two corollaries no longer hold if we replace the last narrowing step, which uses the ϵ-rule, by a simple unification of the left and the right hand side of G_n. Using the ϵ-rule requires not only that the left and the right hand side are unifiable but also that none of the tests detects a redundancy.

Unfortunately, even if there exists a minimal set of solutions for a given equation, LSE narrowing enumerates not necessarily such kind of set.

Example 5.11 Consider the rules

$$\begin{array}{lll} \pi_1: & f(z,c) & \rightarrow & a, \\ \pi_2: & g(c) & \rightarrow & c. \end{array}$$

where x, y and z are variables and a, c are constants. Starting with the equation $f(g(x), y) \doteq a$ there are two LSE narrowing derivations:

$$\begin{array}{llll} f(g(x),y) \doteq a & \rightsquigarrow_{[1,\pi_2,\{x\leftarrow c\}]} & f(c,y) \doteq a \quad \rightsquigarrow_{[\epsilon,\pi_1,\{y\leftarrow c\}]} & a \doteq a \\ f(g(x),y) \doteq a & \rightsquigarrow_{[\epsilon,\pi_1,\{y\leftarrow c\}]} & a \doteq a & \end{array}$$

$\{\{y \leftarrow c\}\}$ is a minimal set of solutions for the given equation, but LSE narrowing computes the non minimal set of solutions $\{\{x \leftarrow c, y \leftarrow c\}, \{y \leftarrow c\}\}$.

Finally, let us mention how LSE narrowing is related to SL left-to-right basic narrowing.

Proposition 5.12 *Any LSE narrowing derivation is SL left-to-right basic.*

Proof: Let

$$G_0 \quad \rightsquigarrow_{[\upsilon_1,\pi_1,\delta_1]} \quad G_1 \quad \rightsquigarrow_{[\upsilon_2,\pi_2,\delta_2]} \quad \ldots \quad G_{n-1} \quad \rightsquigarrow_{[\upsilon_n,\pi_n,\delta_n]} \quad G_n$$

be a LSE narrowing derivation and

$$H_0 \quad \rightarrow_{[\upsilon_1,\pi_1]} \quad H_1 \quad \rightarrow_{[\upsilon_2,\pi_2]} \ldots \rightarrow_{[\upsilon_n,\pi_n]} \quad H_n$$

with $H_i \overset{\text{def}}{=} \lambda_{i,n}(G_i)$, for $i = 0, \ldots, n$, the corresponding rewriting derivation.

By Proposition 5.9, $\delta_n \circ \ldots \circ \delta_1|_{Var(G_0)}$ is normalized. Thus, $U_0 \stackrel{\text{def}}{=} FuOcc(G_0)$ is sufficiently large on $H_0 = \lambda_{0,n}(G_0)$. By Proposition 5.6, the rewriting derivation is a left reduction. By Lemma 4.6, $U_i \stackrel{\text{def}}{=} LRB(U_{i-1}, \upsilon_i, \pi_i)$ is sufficiently large on H_i for $1 \leq i \leq n$. Thus, $\upsilon_i \in U_{i-1}$, for $i = 1, \ldots, n$.

By induction we get $LRB_i = U_i$ for $0 \leq i \leq n$. Hence, $\upsilon_i \in LRB_i$ for $i = 1, \ldots, n$. Furthermore, since $H_i = \lambda_{i,n}(G_i)$, LRB_i is sufficiently large on G_i for $i = 1, \ldots, n$. □

From Theorem 5.5 we get immediately the following corollaries.

Corollary 5.13 (Hullot 80) *Basic narrowing is complete*

Corollary 5.14 (Herold 86) *Left-to-right basic narrowing is complete.*

Corollary 5.15 (Réty 87) *SL left-to-right basic narrowing is complete.*

6 Normalizing LSE Narrowing

One of the most important optimizations of naive narrowing is normalizing narrowing: after every narrowing step the goal is normalized with respect to the given canonical term rewriting system. This allows us to take advantage of the special properties of rewriting steps compared to narrowing steps. Rewriting steps are special narrowing steps which leave invariant the solution space of the current system of equations and thus do not contribute to the construction of a solution. Naive narrowing does not distinguish rewriting and narrowing steps. Every rewriting step leads to a new path in the search space ("don't know indeterminism"), whereas in a canonical term rewriting system the rewriting steps may be executed in an arbitrary ordering ("don't care indeterminism").

6.1 Normalizing Narrowing

Definition 6.1 (Normalizing Narrowing) Let G be a normalized system of equations. A *normalizing narrowing step*

$$G \quad \rightsquigarrow^{\downarrow}_{[\upsilon,l\to r,\delta]} \quad G'{\downarrow}$$

is given by a narrowing step $G \rightsquigarrow_{[\upsilon,l\to r,\delta]} G'$ followed by a normalization $G' \stackrel{*}{\to}_R G'{\downarrow}$ with $G'{\downarrow}$ normalized.

Since G and $G{\downarrow}$ have the same R-unifiers we may assume that G is already in normal from. Note that $\rightsquigarrow^{\downarrow} \subseteq \rightsquigarrow^{*}$ but in general $\rightsquigarrow^{\downarrow} \not\subseteq \rightsquigarrow$.

It is not possible to associate with each rewriting derivation a corresponding *normalizing* narrowing derivation where the same rules are applied at the same occurrences. However, for any rewriting derivation $\sigma(G) \stackrel{*}{\to} \sigma(G){\downarrow}$, where σ is normalized and $\sigma(G){\downarrow}$ is in normal form, there exists another rewriting derivation $\sigma(G) \stackrel{*}{\to} \sigma(G){\downarrow}$ which has a corresponding normalizing narrowing derivation. Moreover, we can assume that the rewriting steps on $\sigma(G)$ corresponding to narrowing steps on G are left reduction steps. This will be used in the proof of the completeness and minimality of normalizing LSE narrowing.

Theorem 6.2 *Consider a normalized system of equations G, a normalized substitution μ and a set of variables V such that $Var(G) \cup Dom(\mu) \subseteq V$. Then there exists a normalization of $H \stackrel{\text{def}}{=} \mu(G)$*

$$
\begin{array}{lllllll}
H = H'_0 & \rightarrow_{[\upsilon_1, l_1 \to r_1]} & H_1 & \rightarrow_{[\upsilon_{11}, l_{11} \to r_{11}]} & \cdots & \rightarrow_{[\upsilon_{1k_1}, l_{1k_1} \to r_{1k_1}]} & H'_1 \\
\vdots & & & & & & \\
H'_{n-1} & \rightarrow_{[\upsilon_n, l_n \to r_n]} & H_n & \rightarrow_{[\upsilon_{n1}, l_{n1} \to r_{n1}]} & \cdots & \rightarrow_{[\upsilon_{nk_n}, l_{nk_n} \to r_{nk_n}]} & H'_n = H{\downarrow},
\end{array}
$$

with left reduction steps $H'_i \rightarrow_{[\upsilon_{i+1}, l_{i+1} \to r_{i+1}]} H_{i+1}$, $i = 0, \ldots, n-1$, such that there exists a normalizing narrowing derivation

$$
\begin{array}{lllllll}
G = G_0{\downarrow} & \leadsto_{[\upsilon_1, l_1 \to r_1, \delta_1]} & G_1 & \rightarrow_{[\upsilon_{11}, l_{11} \to r_{11}]} & \cdots & \rightarrow_{[\upsilon_{1k_1}, l_{1k_1} \to r_{1k_1}]} & G_1{\downarrow} \\
\vdots & & & & & & \\
G_{n-1}{\downarrow} & \leadsto_{[\upsilon_n, l_n \to r_n, \delta_n]} & G_n & \rightarrow_{[\upsilon_{n1}, l_{n1} \to r_{n1}]} & \cdots & \rightarrow_{[\upsilon_{nk_n}, l_{nk_n} \to r_{nk_n}]} & G_n{\downarrow}
\end{array}
$$

which uses the same rules at the same occurrences. Moreover, there exists a normalized substitution λ such that

- $\lambda \circ \delta_n \circ \ldots \circ \delta_1 = \mu \; [V]$
- $H_i = (\lambda \circ \delta_n \circ \ldots \circ \delta_{i+1})\,(G_i), i = 1, \ldots, n$
- $H'_i = (\lambda \circ \delta_n \circ \ldots \circ \delta_{i+1})\,(G_i{\downarrow}), i = 0, \ldots, n$
- $\lambda(G_n{\downarrow}) = H{\downarrow}$.

Proof: By noetherian induction on the rewriting relation $\rightarrow$.

If $H = \mu(G)$ is in normal form, then G is also in normal from and the proposition holds trivially with $\lambda = \mu$.

If H is not in normal form, then there exists a left reduction step $H \rightarrow_{[\upsilon_1, l_1 \to r_1]} H_1$ which by Proposition 3.3 can be lifted to a narrowing step $G \leadsto_{[\upsilon_1, l_1 \to r_1, \delta_1]} G_1$. Moreover, there exists a normalized substitution ψ with $\mu = \psi \circ \delta_1 \; [V]$ and $H_1 = \psi(G_1)$. If $G_1 \rightarrow_{[\upsilon_{11}, l_{11} \to r_{11}]} \cdots \rightarrow_{[\upsilon_{1k_1}, l_{1k_1} \to r_{1k_1}]} G_1{\downarrow}$ is a normalization of G_1, then by the stability of the rewriting relation under substitutions $\psi(G_1) = H_1 \rightarrow_{[\upsilon_{11}, l_{11} \to r_{11}]} \cdots \rightarrow_{[\upsilon_{1k_1}, l_{1k_1} \to r_{1k_1}]} \psi(G_1{\downarrow}) = H'_1$.

Let $V' \stackrel{\text{def}}{=} V \cup Im(\delta_1)$. By applying the induction hypothesis, we obtain a normalization

$$
\begin{array}{lllllll}
H'_1 & \rightarrow_{[\upsilon_2, l_2 \to r_2]} & H_2 & \rightarrow_{[\upsilon_{21}, l_{21} \to r_{21}]} & \cdots & \rightarrow_{[\upsilon_{2k_2}, l_{2k_2} \to r_{2k_2}]} & H'_2 \\
\vdots & & & & & & \\
H'_{n-1} & \rightarrow_{[\upsilon_n, l_n \to r_n]} & H_n & \rightarrow_{[\upsilon_{n1}, l_{n1} \to r_{n1}]} & \cdots & \rightarrow_{[\upsilon_{nk_n}, l_{nk_n} \to r_{nk_n}]} & H'_n = H{\downarrow},
\end{array}
$$

with left reduction steps $H'_i \rightarrow_{[\upsilon_{i+1}, l_{i+1} \to r_{i+1}]} H_{i+1}$, $i = 1, \ldots, n-1$, and a corresponding narrowing derivation

$$
\begin{array}{lllllll}
G_1{\downarrow} & \leadsto_{[\upsilon_2, l_2 \to r_2, \delta_2]} & G_2 & \rightarrow_{[\upsilon_{21}, l_{21} \to r_{21}]} & \cdots & \rightarrow_{[\upsilon_{2k_2}, l_{2k_2} \to r_{2k_2}]} & G_2{\downarrow} \\
\vdots & & & & & & \\
G_{n-1}{\downarrow} & \leadsto_{[\upsilon_n, l_n \to r_n, \delta_n]} & G_n & \rightarrow_{[\upsilon_{n1}, l_{n1} \to r_{n1}]} & \cdots & \rightarrow_{[\upsilon_{nk_n}, l_{nk_n} \to r_{nk_n}]} & G_n{\downarrow}.
\end{array}
$$

Moreover, there is a substitution λ such that

- $\lambda \circ \delta_n \circ \ldots \circ \delta_2 = \psi \ [V']$
- $H_i = (\lambda \circ \delta_n \circ \ldots \circ \delta_{i+1})\,(G_i), i = 2, \ldots, n$
- $H'_i = (\lambda \circ \delta_n \circ \ldots \circ \delta_{i+1})\,(G_i{\downarrow}), i = 1, \ldots, n$
- $\lambda(G_n{\downarrow}) = H{\downarrow}$.

From $\mu = \psi \circ \delta_1 \ [V]$ and $\psi = \lambda \circ \delta_n \circ \ldots \circ \delta_2 \ [V']$, we get $\lambda \circ \delta_n \circ \ldots \circ \delta_1 = \mu \ [V]$ and the proposition follows. □

Since $\rightsquigarrow^{\downarrow} \subseteq \rightsquigarrow^{*}$, Proposition 3.4 holds also for normalizing narrowing strategies. It follows:

Corollary 6.3 (Réty et al. 85) *Normalizing narrowing is complete.*

6.2 Normalizing LSE Narrowing

Our aim is now to extend the idea of LSE narrowing to the case of normalizing narrowing. We can use essentially the same definition as before. Again, the tests have to be applied to the goals where a narrowing step has taken place. These are the goals $G_i{\downarrow}, i = 0, \ldots, n-1$.

Definition 6.4 (Normalizing LSE Narrowing) In a normalizing narrowing derivation

$$\begin{array}{llll} G_0{\downarrow} & \rightsquigarrow_{[v_1,\pi_1,\delta_1]} & G_1 & \stackrel{*}{\rightarrow} & G_1{\downarrow} \\ \vdots & & & & \\ G_{n-1}{\downarrow} & \rightsquigarrow_{[v_n,\pi_n,\delta_n]} & G_n & \stackrel{*}{\rightarrow} & G_n{\downarrow} \end{array}$$

the step $G_{n-1}{\downarrow} \rightsquigarrow G_n \stackrel{*}{\rightarrow} G_n{\downarrow}$ is called a *LSE step* iff the following three conditions are satisfied

(Left-Test) For all $i \in \{0, \ldots, n-1\}$ the subterms of $\lambda_{i,n}(G_i{\downarrow})$ which lie strictly left of v_{i+1} are in normal form.

(Sub-Test) For all $i \in \{0, \ldots, n-1\}$ the proper subterms of $\lambda_{i,n}(G_i{\downarrow}/v_{i+1})$ are in normal form.

(Epsilon-Test) For all $i \in \{0, \ldots, n-1\}$ the term $\lambda_{i,n}(G_i{\downarrow}/v_{i+1})$ is not reducible at occurrence ϵ with a rule smaller than π_{i+1}.

A normalizing narrowing derivation is called a *normalizing LSE narrowing derivation* iff all steps are LSE steps.

The following proposition extends Proposition 5.4 to the case of normalizing narrowing.

Proposition 6.5 *The normalizing narrowing derivation constructed in Theorem 6.2 is a normalizing LSE derivation.*

Proof: Analogous to the proof of Proposition 5.4 with $G_i{\downarrow}$ instead of G_i. □

As an immediate consequence we get:

Theorem 6.6 *Normalizing LSE narrowing is complete.*

Next we extend Proposition 5.6 to the normalizing case.

Proposition 6.7 *Let*

$$G = G_0{\downarrow} \rightsquigarrow_{[\upsilon_1,\pi_1,\delta_1]} G_1 \xrightarrow{*} G_1{\downarrow} \rightsquigarrow_{[\upsilon_2,\pi_2,\delta_2]} \dots \rightsquigarrow_{[\upsilon_n,\pi_n,\delta_n]} G_n \xrightarrow{*} G_n{\downarrow}$$

be a normalizing LSE narrowing derivation.

Then in the corresponding rewriting derivation

$$H = H_0' \rightarrow_{[\upsilon_1,\pi_1]} H_1 \xrightarrow{*} H_1' \rightarrow_{[\upsilon_2,\pi_2]} \dots \rightarrow_{[\upsilon_n,\pi_n]} H_n \xrightarrow{*} H_n' = H{\downarrow},$$

where $H_i \stackrel{\text{def}}{=} \lambda_{i,n}(G_i)$, *for* $i = 1, \dots, n$ *and* $H_i' \stackrel{\text{def}}{=} \lambda_{i,n}(G_i{\downarrow})$, *for* $i = 0, \dots, n$, *the steps* $H_i' \rightarrow_{[\pi_{i+1},\upsilon_{i+1}]} H_{i+1}$ *are left reduction steps, for all* $i = 0, \dots, n-1$.

Proof: Analogous to the proof of Proposition 5.6 with H_i' instead of H_i and $G_i{\downarrow}$ instead of G_i. □

Using this proposition we are now able to prove the minimality result for normalizing LSE narrowing.

Theorem 6.8 *Consider two normalizing LSE narrowing derivations*

$$G_0 \rightsquigarrow^{\downarrow}_{[\upsilon_1,\pi_1,\delta_1]} G_1{\downarrow} \rightsquigarrow^{\downarrow}_{[\upsilon_2,\pi_2,\delta_2]} \dots \rightsquigarrow^{\downarrow}_{[\upsilon_n,\pi_n,\delta_n]} G_n{\downarrow}$$

$$G_0' \rightsquigarrow^{\downarrow}_{[\upsilon_1',\pi_1',\delta_1']} G_1'{\downarrow} \rightsquigarrow^{\downarrow}_{[\upsilon_2',\pi_2',\delta_2']} \dots \rightsquigarrow^{\downarrow}_{[\upsilon_m',\pi_m',\delta_m']} G_m'{\downarrow}.$$

Let $\lambda_{i,n} \stackrel{\text{def}}{=} \delta_n \circ \dots \circ \delta_{i+1}$, *for* $i = 0, \dots, n$ *and* $\lambda_{j,m}' \stackrel{\text{def}}{=} \delta_m' \circ \dots \circ \delta_{j+1}'$ *for* $j = 0, \dots, m$, *where* $\lambda_{n,n} = \lambda_{m,m}' = id$. *Suppose that* G_0 *and* G_0' *respectively* $\lambda_{0,n}$ *and* $\lambda_{0,m}'$ *are identical up to variable renaming, that is there exist substitutions* $\tau_0, \tau_0', \rho, \rho'$ *such that*

- $\tau_0(G_0) = G_0', \quad \tau_0'(G_0') = G_0$ *and*
- $\lambda_{0,n} = \rho' \circ \lambda_{0,m}' \circ \tau_0\ [Var(G_0)], \quad \lambda_{0,m}' = \rho \circ \lambda_{0,n} \circ \tau_0'\ [Var(G_0')]$

Then the two derivations are identical up to variable renaming, that is

- $n = m$
- $\upsilon_i = \upsilon_i'$, *for* $i = 1, \dots, n$
- $\pi_i = \pi_i'$, *for* $i = 1, \dots, n$
- *there exist substitutions* τ_i, τ_i' *such that*
  - $\tau_i(G_i{\downarrow}) = G_i'{\downarrow}, \quad \tau_i'(G_i'{\downarrow}) = G_i{\downarrow}$,
  - $\lambda_{i,n} = \rho' \circ \lambda_{i,m}' \circ \tau_i\ [Var(G_i)], \quad \lambda_{i,m}' = \rho \circ \lambda_{i,n} \circ \tau_i'\ [Var(G_i')]$,

for $i = 1, \ldots, n$.

Proof: Without loss of generality we assume $n \leq m$. First we show by induction on n that the first n steps of the two derivations are identical up to variable renaming. For $n = 0$ nothing has to be shown. Assume therefore $n \geq 1$ and consider the associated rewriting derivations

$$\lambda_{0,n}(G_0) \rightarrow_{[\upsilon_1,\pi_1]} \lambda_{1,n}(G_1) \stackrel{*}{\rightarrow} \lambda_{1,n}(G_1\!\downarrow) \rightarrow_{[\upsilon_2,\pi_2]} \ldots \stackrel{*}{\rightarrow} \lambda_{n,n}(G_n\!\downarrow)$$

and

$$\lambda'_{0,m}(G'_0) \rightarrow_{[\upsilon'_1,\pi'_1]} \lambda'_{1,m}(G'_1) \stackrel{*}{\rightarrow} \lambda'_{1,m}(G'_1\!\downarrow) \rightarrow_{[\upsilon'_2,\pi'_2]} \ldots \stackrel{*}{\rightarrow} \lambda'_{m,m}(G'_m\!\downarrow).$$

By Proposition 6.7, the rewriting steps $\lambda_{0,n}(G_0) \rightarrow_{[\upsilon_1,\pi_1]} \lambda_{1,n}(G_1)$ and $\lambda'_{0,m}(G'_0) \rightarrow_{[\upsilon'_1,\pi'_1]} \lambda'_{1,m}(G'_1)$ are both left reduction steps. From $\lambda'_{0,m}(G'_0) = (\rho \circ \lambda_{0,n} \circ \tau'_0)(G'_0) = \rho(\lambda_{0,n}(G_0))$ and $\lambda_{0,n}(G_0) = (\rho' \circ \lambda'_{0,m} \circ \tau_0)(G_0)) = \rho'(\lambda'_{0,m}(G'_0))$ we deduce that $\lambda_{0,n}(G_0)$ and $\lambda'_{0,m}(G'_0)$ are identical up to variable renaming. By the unicity of left reductions, this implies $\upsilon_1 = \upsilon'_1$ and $\pi_1 = \pi'_1$.

From Lemma 3.7 and its proof we get the existence of substitutions τ_1 and τ'_1 with $\tau_1 \circ \delta_1 = \delta'_1 \circ \tau_0$ and $\tau'_1 \circ \delta'_1 = \delta_1 \circ \tau'_0$ such that $\tau_1(G_1) = G'_1$ and $\tau'_1(G'_1) = G_1$. Since τ_1 and τ'_1 are renaming substitutions, we get even $\tau_1(G_1\!\downarrow) = \tau_1(G_1)\!\downarrow = G'_1\!\downarrow$ and $\tau'_1(G'_1\!\downarrow) = \tau'_1(G'_1)\!\downarrow = G_1\!\downarrow$.

From $\lambda_{1,n} \circ \delta_1 = \lambda_{0,n} = \rho' \circ \lambda'_{0,m} \circ \tau_0 = \rho' \circ \lambda'_{1,m} \circ \delta'_1 \circ \tau_0 = \rho' \circ \lambda'_{1,m} \circ \tau_1 \circ \delta_1$ $[Var(G_0)]$ we deduce $\lambda_{1,n} = \rho' \circ \lambda'_{1,m} \circ \tau_1$ $[(Var(G_0) \setminus Dom(\delta_1)) \cup Im(\delta_1|_{Var(G_0)})]$. Since $(Var(G_0) \setminus Dom(\delta_1)) \cup Im(\delta_1|_{Var(G_0)}) = Var(\delta_1(G_0)) = Var(\delta_1(G_0[\upsilon_1 \leftarrow l_1])) \supseteq Var(\delta_1(G_0[\upsilon_1 \leftarrow r_1])) = Var(G_1) \supseteq Var(G_1\!\downarrow)$, this implies $\lambda_{1,n} = \rho' \circ \lambda'_{1,m} \circ \tau_1$ $[Var(G_1\!\downarrow)]$. In the same way, we can show that $\lambda'_{1,m} = \rho \circ \lambda_{1,n} \circ \tau'_1$ $[Var(G'_1\!\downarrow)]$.

Now we can apply the induction hypothesis and we get

- $\upsilon_i = \upsilon_i$, for $i = 2, \ldots, n$
- $\pi_i = \pi'_i$, for $i = 2, \ldots, n$
- there exist substitutions τ_i, τ'_i such that
  - $\tau_i(G_i\!\downarrow) = G'_i\!\downarrow, \quad \tau'_i(G'_i\!\downarrow) = G_i\!\downarrow,$
  - $\lambda_{i,n} = \rho' \circ \lambda'_{i,m} \circ \tau_i\ [Var(G_i)], \quad \lambda'_{i,m} = \rho \circ \lambda_{i,n} \circ \tau'_i\ [Var(G'_i)],$

  for $i = 2, \ldots, n$.

Finally, let us show that $n = m$. Assume $n < m$ holds and consider the derivation $G'_n\!\downarrow \leadsto_{[\upsilon'_{n+1},\pi'_{n+1},\delta'_{n+1}]} G'_{n+1} \stackrel{*}{\rightarrow} G'_{n+1}\!\downarrow \leadsto^{\downarrow} \ldots \leadsto^{\downarrow} G'_m\!\downarrow$. Then $\lambda'_{n,m}(G'_n\!\downarrow)$ is reducible at occurrence υ'_{n+1} with rule π'_{n+1}. Since $G_n\!\downarrow = \lambda_{n,n}(G_n\!\downarrow) = (\rho' \circ \lambda'_{n,m} \circ \tau_n)(G_n\!\downarrow) = \rho'(\lambda'_{n,m}(G'_n\!\downarrow))$ and $\rightarrow$ is stable under substitutions, this would imply that $G_n\!\downarrow$ is reducible in contradiction to the fact that $G_n\!\downarrow$ is the normal form of G_n. This shows that $n = m$ and the theorem is proved. $\square$

Assuming again that narrowing derivations starting from the same goal and using the same rules at the same occurrences produce the same narrowing substitution we get:

Corollary 6.9 *If normalizing LSE narrowing enumerates two solutions σ and σ' which coincide up to variable renaming, then $\sigma = \sigma'$ holds and the two derivations coincide.*

Theorem 6.10 *For any normalizing LSE narrowing derivation*

$$G_0\!\downarrow \rightsquigarrow_{[v_1,\pi_1,\delta_1]} G_1 \xrightarrow{*} G_1\!\downarrow \rightsquigarrow_{[v_2,\pi_2,\delta_2]} \ldots \rightsquigarrow_{[v_n\pi_n,\delta_n]} G_n \xrightarrow{*} G_n\!\downarrow$$

the narrowing substitution $\delta_n \circ \ldots \circ \delta_1|_{Var(G_0\downarrow)}$ is normalized.

Proof: Analogous to the proof of Proposition 5.9 with $G_i\!\downarrow$ instead of G_i. □

Corollary 6.11 *Normalizing LSE narrowing enumerates only normalized substitutions.*

Example 5.11 is still valid if we use normalizing LSE narrowing instead of LSE narrowing.

6.3 Normalizing LSE and SL Left-to-Right Basic Narrowing

Finally, we want to investigate the relationship of normalizing LSE narrowing to normalizing left-to-right basic narrowing as studied in [Rét87, Rét88]. It is well-known that a naive combination of (left-to-right) basic narrowing and normalizing narrowing is not complete.

For rewriting derivations the computation of the sets of basic occurrences is more complicated than for narrowing derivations. We need the notion of *weakly basic rewriting derivation* [Rét87].

Definition 6.12 (Antecedent) Let $t \rightarrow_{[v,l\rightarrow r]} t'$ be a rewriting step. We say that the occurrence ω in t is an *antecedent* of the occurrence ω' in t' iff

- $\omega = \omega'$ and neither $\omega \preceq v$ nor $v \preceq \omega$ or
- there exists an occurrence ρ' of a variable x in r such that $\omega' = v.\rho'.o$ and $\omega = v.\rho.o$ where ρ is an occurrence of the same variable x in l.

See Fig. 2 for illustration.

Definition 6.13 (Weakly Basic Rewriting) Given a rewriting derivation

$$G_1 \quad \rightarrow_{[v_1,l_1\rightarrow r_1]} \quad G_2 \quad \rightarrow_{[v_2,l_2\rightarrow r_2]} \ldots \rightarrow_{[v_{n-1},l_{n-1}\rightarrow r_{n-1}]} \quad G_n$$

and a set $WB_1 \subseteq Occ(G_1)$ of occurrences in G_1 the corresponding sets of *weakly basic occurrences* are inductively defined by

$$\begin{aligned} WB_{i+1} \overset{\text{def}}{=} \; & (WB_i \setminus \{v \in WB_i \mid v \succeq v_i\}) \cup \{v_i.o \mid o \in FuOcc(r_i)\} \cup \{v \in Occ(G_{i+1}) \mid \\ & v = v_i.o, o \notin FuOcc(r_i) \text{ and all antecedents of } v \text{ in } G_i \text{ are in } WB_i\}, \end{aligned}$$

for $i = 1, \ldots, n-1$. The rewriting derivation is *weakly based on* WB_1 iff $v_i \in WB_i$, for all $i = 1, \ldots, n-1$. Instead of WB_n we will also write $WB(WB_1, G_1 \xrightarrow{*} G_n)$.

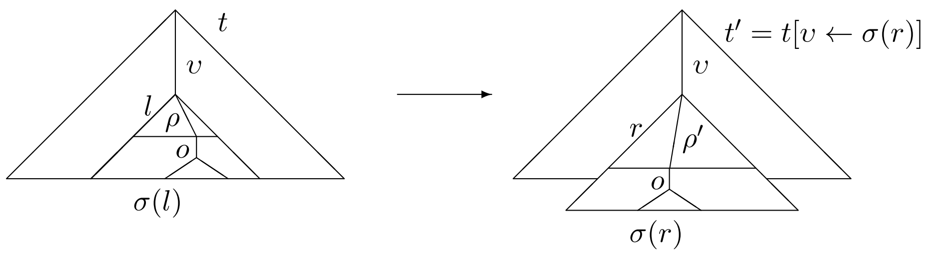

Figure 2: Illustration of antecedents: the occurrence $\upsilon.\rho.o$ is an antecedent of $\upsilon.\rho'.o$.

The main difference compared to the computation of the set B_{i+1} of basic occurrences is that occurrences under υ_i which do not correspond to non-variable occurrences in r_i may belong to WB_{i+1}. Note that different reductions $G \xrightarrow{*} G'$ can lead to different sets $WB(U, G \xrightarrow{*} G')$. But this does not affect the completeness of the narrowing strategies introduced below (see [Wer91]).

The notion of weakly basic occurrences is closely related to the notion of sufficient largeness as is illustrated by the following lemma [Rét87].

Lemma 6.14 *Let U be sufficiently large on G_1. Then any rewriting derivation $G_1 \xrightarrow{*} G_n$ is weakly based on U, and $WB(U, G_1 \xrightarrow{*} G_n)$ is sufficiently large on G_n.*

Proof: By induction on the length of the derivation. In the case $n = 1$ nothing has to be shown. Assume therefore that $G_1 \xrightarrow{*} G_{n-1}, n > 1$, is weakly based on $U = WB_1$ and that WB_{n-1} is sufficiently large on G_{n-1}. Then the occurrence υ_{n-1} in $G_{n-1} \rightarrow_{[\upsilon_{n-1}, \pi_{n-1}]} G_n$ must belong to WB_{n-1}. If $\omega_n \in Occ(G_n) \setminus WB_n$, then at least one antecedent ω_{n-1} of ω_n in G_{n-1} does not belong to WB_{n-1}. Since WB_{n-1} is sufficiently large on G_{n-1}, we deduce that $G_n/\omega_n = G_{n-1}/\omega_{n-1}$ is irreducible. This shows that WB_n is sufficiently large on G_n. □

Definition 6.15 (Normalizing SL Left-to-Right Basic Narrowing) Let $G_0{\downarrow}$ be a normalized system of equations. A derivation

$$
\begin{array}{llllll}
(G_0{\downarrow}, U_0{\downarrow}) & \rightsquigarrow_{[\upsilon_1, \pi_1, \delta_1]} & (G_1, U_1) & \xrightarrow{*} & (G_1{\downarrow}, U_1{\downarrow}) \\
\vdots & & & & \\
(G_{n-1}{\downarrow}, U_{n-1}{\downarrow}) & \rightsquigarrow_{[\upsilon_n, \pi_n, \delta_n]} & (G_n, U_n) & \xrightarrow{*} & (G_n{\downarrow}, U_n{\downarrow}),
\end{array}
$$

with $U_0{\downarrow} \stackrel{\text{def}}{=} FuOcc(G_0{\downarrow})$, $U_i \stackrel{\text{def}}{=} LRB(U_{i-1}{\downarrow}, \upsilon_i, \pi_i)$ and $U_i{\downarrow} \stackrel{\text{def}}{=} WB(U_i, G_i \xrightarrow{*} G_i{\downarrow})$ is called a *normalizing left-to-right basic narrowing derivation* if for $i = 1, \ldots, n$

- $\upsilon_i \in U_{i-1}{\downarrow}$,
- $G_i \xrightarrow{*} G_i{\downarrow}$ is weakly based on U_i, and
- $G_i{\downarrow}$ is normalized.

The derivation is called a *normalizing SL left-to-right basic narrowing derivation* if moreover $U_i{\downarrow}$ is sufficiently large on $G_i{\downarrow}$, for all $i = 1, \ldots, n$.

Theorem 6.16 *Any normalizing LSE narrowing derivation is also a normalizing SL left-to-right basic narrowing derivation.*

Proof: Consider a normalizing LSE narrowing derivation

$$G_0{\downarrow} \rightsquigarrow_{[\upsilon_1,\pi_1,\delta_1]} G_1 \xrightarrow{*} G_1{\downarrow} \rightsquigarrow_{[\upsilon_2,\pi_2,\delta_2]} \cdots \rightsquigarrow_{[\upsilon_n\pi_n,\delta_n]} G_n \xrightarrow{*} G_n{\downarrow}$$

and the corresponding rewriting derivation

$$H = H'_0 \rightarrow_{[\upsilon_1,\pi_1]} H_1 \xrightarrow{*} H'_1 \rightarrow_{[\upsilon_2,\pi_2]} \cdots \rightarrow_{[\upsilon_n,\pi_n]} H_n \xrightarrow{*} H'_n = H{\downarrow},$$

with $H_i \stackrel{\text{def}}{=} \lambda_{i,n}(G_i)$, for $i = 1, \ldots, n$, and $H'_i \stackrel{\text{def}}{=} \lambda_{i,n}(G_i{\downarrow})$, for $i = 0, \ldots, n$. Define the sets of occurrences $U_0{\downarrow} \stackrel{\text{def}}{=} \mathit{FuOcc}(G_0{\downarrow})$, $U_i \stackrel{\text{def}}{=} \mathit{LRB}(U_{i-1}{\downarrow}, \upsilon_i, \pi_i)$ and $U_i{\downarrow} \stackrel{\text{def}}{=} \mathit{WB}(U_i, G_i \xrightarrow{*} G_i{\downarrow})$, for $i = 1, \ldots, n$.

By induction on $j = 0, \ldots, n$, we prove for all $i = 1, \ldots, j$ that

- $\upsilon_i \in U_{i-1}{\downarrow}$,
- $G_i \xrightarrow{*} G_i{\downarrow}$ is weakly based on U_i

and for all $i = 0, \ldots, j$ that

- $U_i{\downarrow}$ is sufficiently large on $G_i{\downarrow}$ and H'_i.

Since $\lambda_{0,n}|_{Var(G_0{\downarrow})}$ is normalized by Theorem 6.10, $U_0{\downarrow} = \mathit{FuOcc}(G_0{\downarrow})$ is sufficiently large on both $G_0{\downarrow}$ and $H'_0 = \lambda_{0,n}(G_0{\downarrow})$.

Suppose the statement is true for $0 \leq j - 1 < n$. By the induction hypothesis the statement holds for all $i < j$. Hence, $U_{j-1}{\downarrow}$ is sufficently large on H'_{j-1}. Consider the rewriting step $H'_{j-1} \rightarrow_{[\upsilon_j,\pi_j]} H_j$. Since $U_{j-1}{\downarrow}$ is sufficiently large on H'_{j-1}, we get $\upsilon_j \in U_{j-1}{\downarrow}$. By Proposition 6.7, $H'_{i-1} \rightarrow_{[\upsilon_i,\pi_i]} H_i$ is a left reduction step. Using Lemma 4.6, we can conclude that the set $U_j = \mathit{LRB}(U_{j-1}{\downarrow}, \upsilon_j, \pi_j)$ is sufficiently large on H_j. Since $H_j = \lambda_{j,n}(G_j)$, this shows that U_j is also sufficiently large on G_j. By Lemma 6.14, this implies that the rewriting derivations $G_j \xrightarrow{*} G_j{\downarrow}$ and $H_j \xrightarrow{*} H'_j$ are weakly based on U_j and that $U_j{\downarrow} = \mathit{WB}(U_j, G_i \xrightarrow{*} G_j{\downarrow}) = \mathit{WB}(U_j, H_i \xrightarrow{*} H'_j)$ is sufficiently large on $G_j{\downarrow}$ and H'_j. Therefore, the statement is true for j.

□

Corollary 6.17 *Normalizing SL left-to-right basic narrowing is complete.*

7 Empirical Results

In this last section, we give a number of empirical results to illustrate the various narrowing strategies. In particular we show how the narrowing search space can be reduced using the LSE strategy.

Our computations have been done in the KArlsruhe NArrowing Labor KANAL [Kri90] which is implemented in the Prolog dialect KA-Prolog on a SUN SPARC 10/41.

We will proceed in two steps. First we give for a very simple example the size of the narrowing tree for all strategies which have been discussed in this paper. In this example, LSE narrowing yields the same results as SL left-to-right basic narrowing.

Then we focus on the most efficient strategies for arbitrary canonical systems, namely normalizing left-to-right basic narrowing, normalizing SL left-to-right basic narrowing and normalizing LSE narrowing and show on some larger examples how the narrowing search space can be reduced by the various reducibility tests.

7.1 Comparing all narrowing strategies for a functional term rewrite system

Consider the canonical term rewriting system

$$\begin{array}{rcll} R & = & \{ & 0+x \to x, \quad s(x)+y \to s(x+y), \\ & & & 0*x \to 0, \quad s(x)*y \to y+x*y \quad \} \end{array}$$

for the addition and multiplication of natural numbers. This term rewrite system is *functional* in the sense of [DG89]: the rules are constructor-based, left-linear and non-overlapping.

We would like to answer the query

$$?- \quad x*x+y*y \doteq s(0)$$

which has two solutions

$$\sigma_1 = \{x \leftarrow 0, y \leftarrow s(0)\} \text{ and } \sigma_2 = \{x \leftarrow s(0), y \leftarrow 0\}.$$

First we consider the narrowing strategies without normalization. The solution σ_1 is found in depth 6, the solution σ_2 in depth 7 of the narrowing tree.

The number of nodes in the narrowing tree is given in Fig. 3. The numbers for LSE narrowing are the same as for SL left-to-right basic narrowing.

If we do narrowing with normalization both solutions are found in depth 3 and many fewer narrowing steps are needed. The naive narrowing tree contains 51372 nodes at depth 7 whereas in the normal tree at depth 3 there are only 72. Although normalizing narrowing steps are more costly than naive narrowing steps, this is an enormous gain of efficiency (see Fig. 4). Again there is no difference between normalizing SL left-to-right basic narrowing and normalizing LSE narrowing.

7.2 Comparing the best narrowing strategies for a arbitrary canonical system

For simple term rewrite systems, LSE narrowing does not improve the performance of narrowing compared to other strategies. However, with increasing

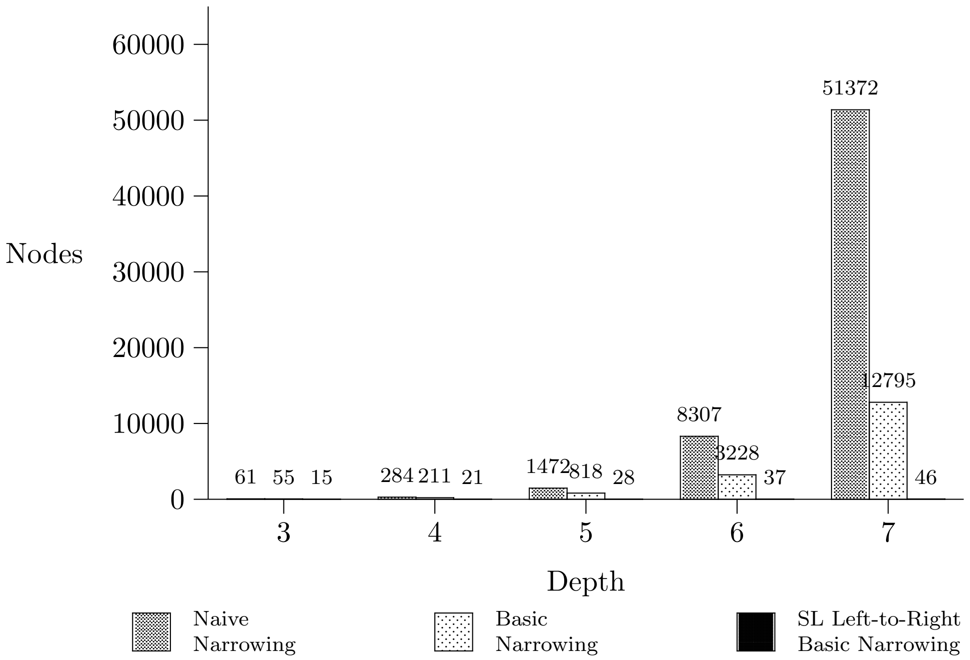

Figure 3: Non-normalizing strategies

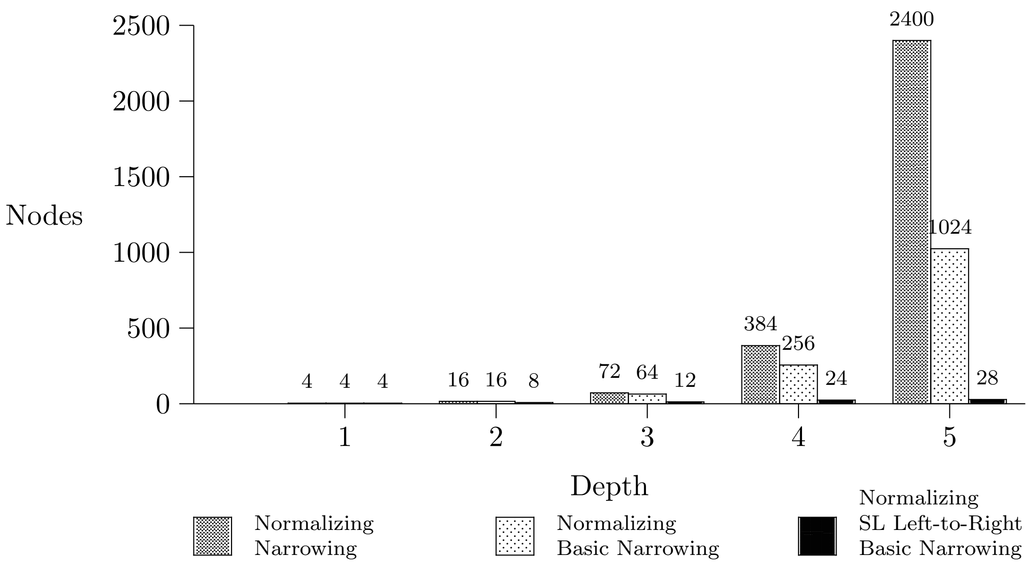

Figure 4: Normalizing strategies

complexity of the rewrite systems and queries, the LSE strategy becomes more and more important. In particular, the following properties of rewriting systems and goals are relevant for LSE narrowing:

- overlapping left-hand sides
- non-regular rules $l \to r$
- left-hand sides with several defined function symbols
- non-linear rule sides and non-linear goals

To illustrate these points, consider a family R_n of canonical rewriting systems for arithmetic modulo an integer number $n \geq 1$.

$$R_n = \{ \quad \begin{array}{rcl} 0 + x & \to & x \\ s(x) + y & \to & s(x+y) \\ x + s(y) & \to & s(x+y) \\ 0 * x & \to & 0 \\ s(x) * y & \to & x * y + y \\ \underbrace{s(s(\ldots s}_{\text{n-times}}(x)\ldots)) & \to & x \\ (\ldots((x + \underbrace{y) + y) + \ldots y)}_{\text{n-times}} & \to & x \\ (\ldots((\underbrace{x + x) + x) + \ldots + x)}_{\text{n-times}} & \to & 0 \quad \}. \end{array}$$

With increasing n, the last three rules generate more and more redundancies in the normalizing SL left-to-right basic narrowing tree which can be eliminated by the LSE-Tests.

We solve the goal

$$?- \quad x * y + x \doteq s(0).$$

for the systems $R_3, R_5, R_7, R_9, R_{11}, R_{13}$. For the system R_n the narrowing tree has depth $2n+1$, then no more derivations are possible. If we compare the running time (in seconds) needed by normalizing SL left-to-right basic narrowing and LSE narrowing in order to compute the narrowing tree of depth $2n+1$ for the system R_n, we get the following results.

System	Depth	Time: SL	Time: LSE	Time: Factor
R_3	7	0	0	1
R_5	11	4	3	1,3
R_7	15	43	15	2,8
R_9	19	383	49	7,8
R_{11}	23	3366	142	23,7
R_{13}	27	23387	307	76,1

Note that in this example the performance of LSE narrowing increases although the derivations get longer and longer.

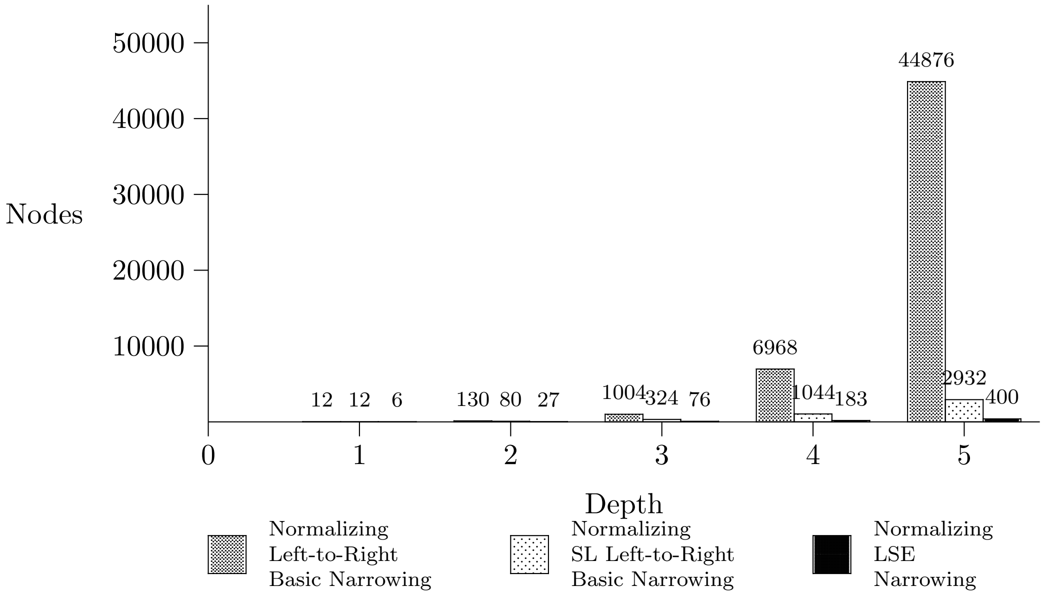

Figure 5: Best narrowing strategies for arbitrary systems (only normalized substitutions)

We finish this section with a large practical example where the three tests work together in a very nice way. We consider the canonical term rewriting system for the integer arithmetic given in [RKKL85]

$$
\begin{array}{lllllll}
Int = \{ & s(p(x)) & \to & x, & & & \\
 & p(s(x)) & \to & x, & & & \\
 & 0 + x & \to & x, & x + 0 & \to & x, \\
 & s(x) + y & \to & s(x+y), & x + s(y) & \to & s(x+y), \\
 & p(x) + y & \to & p(x+y), & x + p(y) & \to & p(x+y), \\
 & -0 & \to & 0, & & & \\
 & -s(x) & \to & p(-x), & & & \\
 & -p(x) & \to & s(-x), & & & \\
 & 0 * x & \to & 0, & x * 0 & \to & 0, \\
 & s(x) * y & \to & y + (x * y), & x * s(y) & \to & (x * y) + x, \\
 & p(x) * y & \to & (-y) + (x * y), & x * p(y) & \to & (x * y) + (-x), \\
 & -(-x) & \to & x, & & & \\
 & (-x) + x & \to & 0, & x + (-x) & \to & 0, \\
 & x + ((-x) + z) & \to & z, & (-x) + (x + z) & \to & z, \\
 & -(x+y) & \to & (-y) + (-x), & & & \\
 & (x+y) + z & \to & x + (y+z) \quad \} & & &
\end{array}
$$

and take the goal ?− $x * x + y * y \doteq s^5(0)$.

The number of nodes in the narrowing tree is given in Fig. 5.

To compute the narrowing tree of depth 5, we needed 381 sec. for ordinary normalizing left-to-right basic narrowing, 74 sec. for SL normalizing left-to-right basic narrowing, and 19 sec. for normalizing LSE narrowing.

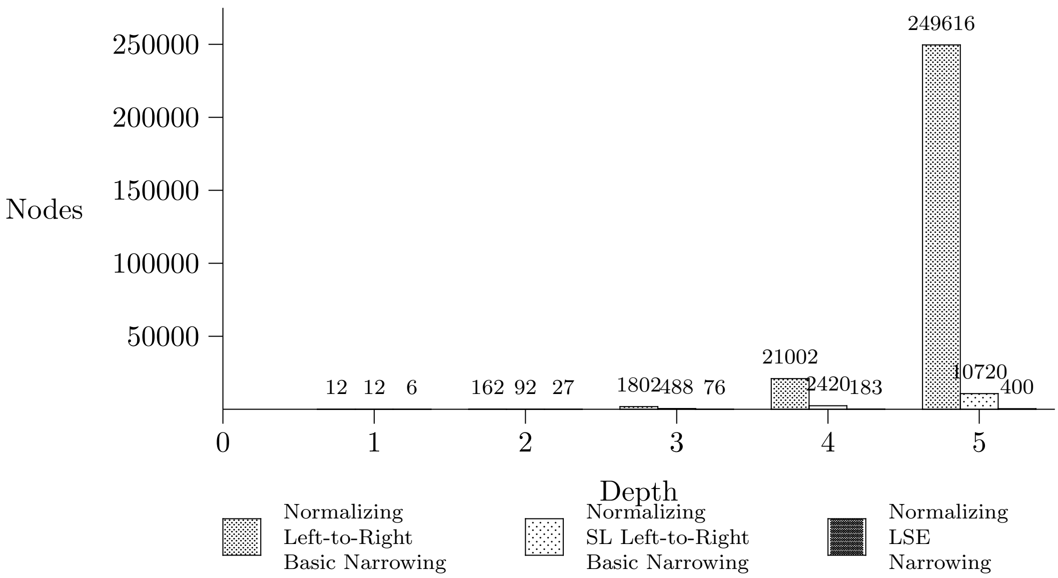

Figure 6: Best narrowing strategies for arbitrary systems (all substitutions)

In all the previous examples, we computed for normalizing left-to-right basic narrowing and normalizing SL left-to-right basic narrowing only narrowing derivations that generate *normalized* narrowing substitutions. For LSE narrowing, this is automatically the case. For the other strategies, however, this makes a big difference: If we admit also non-normalized narrowing substitutions, we get the numbers given in Fig. 6.

These examples illustrate that the reducibility tests done after a narrowing step are just as important for the efficiency of the narrowing procedure as is the choice of the right narrowing strategy.

8 Conclusion

In this paper, we have introduced a new narrowing strategy *LSE narrowing* and its normalizing variant. The main features of LSE narrowing are the following

- there is a one-to-one correspondence between LSE narrowing derivations and left reductions.
- LSE narrowing is complete for arbitrary canonical systems.
- two different LSE narrowing derivations cannot generate the same narrowing substitution.
- LSE narrowing generates only normalized narrowing substitutions.

In a subsequent paper, we will show how LSE narrowing can be realized very efficiently by a slight modification of a WAM-based implementation of left-to-right basic narrowing [WBK93]. According to their definition, LSE narrowing steps seem to be very expensive, because a large number of subterms has to be considered. However, using left-to-right basic occurrences this number can be reduced in a drastic way.

References

[BGM88] P. G. Bosco, E. Giovannetti, and C. Moiso. Narrowing vs. SLD-Resolution. *Theoretical Computer Science*, 59:3 –23, 1988.

[BKW92] A. Bockmayr, S. Krischer, and A. Werner. An optimal narrowing strategy for general canonical systems. In *Conditional Term Rewriting Systems, CTRS'92, Pont-à-Mousson, France.* Springer, LNCS 656, 1992.

[Boc86] A. Bockmayr. Narrowing with inductively defined functions. Technical Report 25/86, Fakultät für Informatik, Univ. Karlsruhe, 1986.

[Boc88] A. Bockmayr. Narrowing with built-in theories. In *First Int. Workshop Algebraic and Logic Programming, Gaußig.* Springer, LNCS 343, 1988.

[DG89] J. Darlington and Y. Guo. Narrowing and unification in functional programming - an evaluation mechanism for absolute set abstraction. In *Proc. Rewriting Techniques and Applications, Chapel Hill.* Springer, LNCS 355, 1989.

[Ech88] R. Echahed. On completeness of narrowing strategies. In *Proc. CAAP'88, Nancy.* Springer, LNCS 299, 1988.

[Fay79] M. Fay. First-order unification in an equational theory. In *4th Workshop on Automated Deduction, Austin, Texas*, 1979.

[Fri85] L. Fribourg. Handling function definitions through innermost superposition and rewriting. In *Proc. Rewriting Techniques and Applications, Dijon.* Springer, LNCS 202, 1985.

[Her86] A. Herold. Narrowing techniques applied to idempotent unification. SEKI-Report SR-86-16, Univ. Kaiserslautern, 1986.

[HO80] G. Huet and D. C. Oppen. Equations and rewrite rules, A survey. In R. V. Book, editor, *Formal Language Theory.* Academic Press, 1980.

[Höl89] S. Hölldobler. *Foundations of Equational Logic Programming.* Springer, LNCS 353, 1989.

[Hul80] J. M. Hullot. Canonical forms and unification. In *Proc. 5th Conference on Automated Deduction, Les Arcs.* Springer, LNCS 87, 1980.

[KB91] S. Krischer and A. Bockmayr. Detecting redundant narrowing derivations by the LSE-SL reducibility test. In *Proc. Rewriting Techniques and Applications, Como.* Springer, LNCS 488, 1991.

[Kri90] S. Krischer. Vergleich und Bewertung von Narrowing-Strategien. Diplomarbeit, Fakultät für Informatik, Univ. Karlsruhe, 1990.

[NRS89] W. Nutt, P. Réty, and G. Smolka. Basic narrowing revisited. *Journal of Symbolic Computation*, 7:295–317, 1989.

[Pad88] P. Padawitz. *Computing in Horn Clause Theories*, volume 16 of *EATCS Monographs*. Springer, 1988.

[Rét87] P. Réty. Improving basic narrowing techniques. In *Proc. Rewriting Techniques and Applications, Bordeaux*. Springer, LNCS 256, 1987.

[Rét88] P. Réty. *Méthodes d'Unification par Surréduction*. PhD thesis, Univ. Nancy, 1988.

[RKKL85] P. Réty, C. Kirchner, H. Kirchner, and P. Lescanne. NARROWER: a new algorithm for unification and its application to logic programming. In *Proc. Rewriting Techniques and Applications, Dijon*. Springer, LNCS 202, 1985.

[WBK93] A. Werner, A. Bockmayr, and S. Krischer. How to realize LSE narrowing. Technical Report 6/93, Fakultät für Informatik, Univ. Karlsruhe, 1993.

[Wer91] A. Werner. Termersetzung und Narrowing mit geordneten Sorten. Diplomarbeit, Fakultät für Informatik, Univ. Karlsruhe, 1991.

[You88] J. You. Solving equations in an equational language. In *First Int. Workshop Algebraic and Logic Programming, Gaußig*. Springer, LNCS 343, 1988.

[You91] J. You. Unification modulo an equality theory for equational logic programming. *J. Comp. Syst. Sc.*, 42:54–75, 1991.