On Relative Normalization in Orthogonal Expression Reduction Systems

John Glauert and Zurab Khasidashvili
School of Information Systems, UEA, Norwich NR4 7TJ England
jrwg@sys.uea.ac.uk, zurab@sys.uea.ac.uk*

May 3, 1994

1 Introduction

Since a normalizable term, in a rewriting system, may have an infinite reduction, it is important to have a strategy that enables one to construct normilizing reductions. Such strategies are called *normalizing*. It is well known that the leftmost-outermost strategy is normalizing in the λ-calculus. For orthogonal Term Rewriting Systems (OTRSs), a general normalizing strategy, called the *needed* strategy, was found by Huet and Lévy in [2]. The needed strategy always contracts a *needed* redex – a redex whose residual is to be contracted in any reduction to normal form. Huet and Lévy showed that any term t not in normal form has a needed redex, and that repeated contraction of needed redexes in t leads to its normal form whenever there is one. They also defined the class of *strongly sequential* OTRSs where a needed redex can efficiently be found in any term.

Barendragt et al. [1] generalized the concept of neededness to the λ-calculus. They studied neededness not only w.r.t. normal forms, but also w.r.t. head-normal forms – a redex is *head-needed* if its residuals are contracted in each reduction to a head-normal form. The authors proved correctness of the two needed strategies for computing normal forms and head-normal forms, respectively. In Khasidashvili [4, 6], a similar normalizing strategy, called the *essential* strategy, was defined for the λ-calculus. The strategy contracts *essential* redexes – the redexes that have *descendants* under any reduction. The notion of descendant is a refinement of that of *residual* – the descendant of a contracted redex is its contractum, while it does not have residuals. An advantage of the correctness proof of the essential strategy as in [4] is that its generalization to all Orthogonal Combinatory Reduction Systems (OCRSs) is straightforward. Kennaway and Sleep [3] used a generalization of Lévy's labelling for the λ-calculus [9] to adapt the proof from [1] to the case of OCRSs.

In [10], Maranget introduced a different notion of neededness, where a redex u is needed if it has a residual under any reduction that does not contract the residuals of u. This neededness notion makes sense also for terms that do not have a normal form. It was shown in [6] that a redex is Maranget-needed iff it is essential. In [10], Maranget studied also a strategy that computes a

*This work was supported by the Science and Engineering Council of Great Britain under grant GR/H 41300

('minimal') weak head-normal form of a term in an OTRS. Normalization w.r.t. another interesting set of 'normal forms', that of constructor head-normal forms in constructor OTRSs, is studied by Nöcker [11]. A question arises naturally: what are the properties that a set of terms must possess in order for the neededness theory of Huet and Lévy still to make sense?

The main contribution of this paper is to provide a solution to that question. We introduce the notion of *neededness* w.r.t. a set of reductions Π or a set of terms M so that each existing notion of neededness can be given by specifying Π or M. We find natural conditions on M, called *stability*, that are sufficient and necessary for each term not in M-normal form (i.e., not in M) to have at least one M-needed redex, and repeated contraction of M-needed redexes in a term t to lead to an M-normal form of t whenever there is one. A set M of terms is stable if it is *closed under reduction* and is *closed under unneeded expansion*: for any $e \xrightarrow{u} o$ such that $e \notin M$ and $o \in M$, a residual of u is contracted in any reduction from e to a term in M. We present a counterexample to show that in general the M-needed strategy is not *hypernormalizing* for every stable M.

Our relative neededness notion is based on tracing of *(open) components*, which are occurrences of contexts not containing *any* bound variable, rather than tracing of redexes or subterms. We therefore introduce a notion of *descendant* for components and show that, for any co-initial reductions P and Q, there are converging reductions P' and Q' such that the descendants of any subterm or a component of the initial term under $P+P'$ (the concatenation of P and P') and $Q+Q'$ are the same occurrences. We call a component of a term $t \notin M$ *M-needed* if at least one descendant of it is 'involved' in any M-normalizing reduction; a redex is *M-needed* if so is its pattern. This approach to defining the neededness notion is crucial from technical point of view, because a 'non-standard' situation – when a component under an M-unneeded component is M-needed – may occur for some stable sets M. Also, components of a term in an OCRS enjoy the same 'disjointness' property that subterms of a term in an OTRS possesses: descendants of disjoint components of a term in an OCRS remain disjoint, and this allows for simpler proofs.

2 Preliminaries

Combinatory Reduction Systems (CRSs) have been introduced in Klop [7] to provide a uniform framework for reductions in Term rewriting Systems (TRSs) and reductions with substitutions (referred to also as higher order rewriting) as in the λ-calculus. Several other interesting formalisms have been introduced later. We refer to Klop et al. [8] for further information.

Our syntax, *Expression Reduction Systems (ERS)* (called CRSs in [5]) is closer to the syntax of λ-calculus and First Order Logic. The alphabet contains (object) variables and function and quantifier symbols. Quantifiers may be *partial*, that is, they may bind variables only in some of the arguments. Patterns of left-hand and right-hand sides of rules in ERSs may have arbitrary shape, *object metavariables* must be bound, and right-hand sides may contain any number of, sometimes nested, *metasubstitutions*. For example, the β-rule is written as $\beta : Ap(\lambda aA, B) \to (B/a)A$, where a is an object metavariable, to be instantiated by a variable, and A and B are term-metavariables, to be instantiated by terms. The expression $(B/a)A$ is a metasubstitution, and its instance $(t/x)s$ denotes the result of substitution of the term t for x in s.

Notation We use u, v, w for redexes, t, s, e, o for terms, P, Q for reductions; $+$ denotes concatenation of reductions; P/Q denotes the *residual* of P by Q [9, 2]; we write $t \subseteq s$ if t is a subterm of s; $pat(u)$ denotes the pattern of u; u will also denote the reduction that contracts the redex u.

We call a *component* an occurrence of a context that does not contain any bound variables (even if their binders are in the component). We use the letters used for terms to denote components as well. We write $s \sqsubseteq t$ if s is a component of t.

In Khasidashvili [5] a notion of *descendant* is introduced to trace subterms during reductions in OERSs. The main difference with the notion of *residual* of redexes is that the descendant of a contracted redex is its contractum, while the redex does not have residuals. In the case when the pattern $pat(u)$ of $u \subseteq t$ does not overlap with a component $s \sqsubseteq t$, one can naturally define *u-descendants* of s to be the descendants of the corresponding subterm $o = e[o_1, \ldots, o_n]$ with the 'corresponding' descendants of $o_1, \ldots, o_n$ removed; the definition generalizes by transitivity to all R-reductions that are *external* to s, i.e., do not execute redexes whose patterns overlap with descendants of s (see Definition 4.1.(1)). The descendant relation for components can be defined also in the case of overlap between the pattern of the contracted redex and the traced component, but the definition has many cases and we leave it out; it anyway is not needed to prove the main result of the paper. We also show that OERSs have the following *strict Church-Rosser* property: for any co-initial P and Q, the descendants of any component and any subterm of the initial term under $P + Q/P$ and $Q + P/Q$ are the same occurrences.

3 Neededness, essentiality and unabsorbness

Recall that a redex u in t is *Huet-Lévy-needed* [2] if in each reduction of t to normal form (if any) at least one residual of u is contracted; u is *Maranget-needed* [10] if u has at least one residual under any reduction starting from t that does not contract residuals of u. Further, a subterm s in t is *essential* (written $ES(s,t)$) if s has at least one descendant under any reduction starting from t and *inessential* (written $IE(s,t)$) otherwise [6]. Finally, a subterm $s \subseteq t$ is *unabsorbed* in $P : t \twoheadrightarrow e$ if none of the descendants of s appear in redex-arguments of terms in P, and is *absorbed in P* otherwise; s is *unabsorbed in t* if it is unabsorbed in any reduction starting from t and *absorbed in t* otherwise [6]. (It is easy to see that a redex $u \subseteq t$ is unabsorbed iff u is *external* [2] in t.) Obviously, unabsorbed redexes are essential and are both Lévy- and Maranget-needed. Note that Lévy- and Maranget-neededness coincide for normalizable terms.

We show that every term t not in normal form has an *eraser* – a reduction that erases all inessential subterms of t. It is easy to show that, for example, *fair* reductions are erasers ($P : t \to t_1 \to \ldots$ is *fair* if each redex $u_i \subseteq t_i$ is *secured* in P, i.e., there is $j > i$ such that u_i does not have residuals in t_j). As in the case of OTRSs [6], we then show that Maranget-neededness (for redexes) coincides with our essentiality in OERSs as well. ($\Leftarrow$) is immediate. For ($\Rightarrow$), we first show that if one replaces any number of inessential subterms $s_1, \ldots, s_n$ in a term s with any terms $t_1, \ldots, t_n$, then $t_1, \ldots, t_n$ will be inessential in the obtained term t — one can construct an eraser of $t_1, \ldots, t_n$ by a kind of simulation of an eraser of $s_1, \ldots, s_n$. Now, if $IE(u,t)$, then any subterm o not containing symbols occurring in redex-patterns is inessential when substituted for u in t, and an eraser of u that does not contract the residuals of u can be constructed from any eraser of o; hence u is not Maranget-needed. A side result is the existance of an essential redex in every term not in normal form. Existence of an unabsorbed redex in any term t not in normal form can be proved exactly as in the case of OTRSs [6].

4 Relative notions of neededness

Definition 4.1 (1) We call a reduction P starting from a term t in an OERS R *external* to a component $e \sqsubseteq t$ if there is no redex executed in P whose pattern overlaps with a descendant of e.

(2) We call a component $e \sqsubseteq t$ *P-unneeded* if P is external to e and call it *P-needed* otherwise.

(3) Let Π be a set of reductions. We call $e \sqsubseteq t$ *Π-needed* if it is P-needed for every $P \in \Pi$ starting from t, and call Π-*unneeded* otherwise.

(4) Let M be a set of terms in R, and let Π_M be the set of *M-normalizing* reductions, i.e., reductions that end at a term in M. We call $e \sqsubseteq t$ *M-needed* (*M-unneeded*) if it is *Π_M-needed* (*Π_M-unneeded*). We say that P *M-suppresses* e if P is M-normalizing and is external to e.

(5) If $o \subseteq t$, then we call o (P, Π, M-)*unneeded* (*needed*) if so is $Int(o)$, the component obtained from the subterm o by removing *all* bound variables. However, if not otherwise stated, we say that a redex $u \subseteq t$ (which is a subterm) is *M-unneeded* (*M-needed*) if so is the pattern $pat(u)$ of u.

(6) We call a set M of terms *stable* if: (a) M is closed under reductions in R; and (b) M is *closed under unneeded expansion*: for any $e \xrightarrow{u} o$ such that $e \notin M$ and $o \in M$, u is M-needed.

We write $NE_{(P,\Pi,M)}(e,t)$ ($UN(e,t)$) if e is (P, Π, M)-needed (unneeded). Below M will always denote a stable set of terms.

If $o \sqsubseteq t$ and $P : t \twoheadrightarrow s$ is external to o, then we call the descendants of o also the *residuals* of o. A component that does overlap with the pattern of a contracted redex does not have residuals. Note that if $u \subseteq t$, then $P : t \twoheadrightarrow s$ is external to $pat(u)$ iff P does not contract the residuals of u, because orthogonality of the system implies that if the pattern $pat(v)$ of a redex v contracted in P does overlap with a residual of $pat(u)$, then $pat(u) = pat(v)$. Hence u is M-needed iff at least one residual of it is contracted in each reduction from t to a term in M (the intended notion of neededness).

The most appealing examples of stable sets for an OERS are the set of normal forms [2], the set of head-normal forms [1], and the set of constructor-head-normal forms for constructor TRSs [11]. The graph G_s of a term s (which consists of terms to which s is reducible) is closed under reduction, but need not be closed under unneeded expansion. For example, the graph $G_{I(x)} = \{I(x), x\}$ of $I(x)$ is closed under reduction but need not be closed under unneeded expansion: $I(I(x))$ can be reduced to $I(x)$ by reducing either I-redex. Note that any set closed under reduction is the union of the graphs of a set of terms. The set M' of strongly normalizable terms (in some OERS) is closed under reduction, but need not be closed under unneeded expansion. For example, $t = Kz(Ky\omega)$ with $w = (\lambda x.xx)(\lambda x.xx)$ and $K = \lambda xy.x$ can be reduced to a strongly normalizable term (in this case, a variable) by contracting either K-redex. Note that *none* of the redexes in t are M'-needed.

Note that Huet-Lévy-neededness (for redexes) coincides with neededness w.r.t. all normalizing reductions. We show that Maranget-neededness is neededness w.r.t. all fair reductions. Indeed, ($\Rightarrow$) is immediate: if u is Maranget-needed in t, then any fair P starting from t should contract a residual of u to secure it. ($\Leftarrow$) follows from the fact that, if u is not Maranget-needed, then any reduction that secures u without contracting its residuals can be extended to a fair one.

5 The Uniform Normalization Theorem

As for the case of subterms in OTRSs, descendants of non-overlapping components are non-overlapping components; due to this property everything works nicely. For example, one can show that if P M-suppresses $pat(u_1), \ldots, pat(u_n) \sqsubseteq t$ and $t \xrightarrow{u} e$, then P/u M-suppresses the u-residuals of $pat(u_1), \ldots, pat(u_n)$. It follows that

(α) residuals of M-unneeded redexes remain M-unneeded.

The fact that

(γ) each M-needed redex, if not contracted, has at least one M-needed residual

is not immediate, but can be proved, e.g., using the main theorem. One can show that for M-unneeded/M-needed components properties similar to the properties of inessential/essential subterms hold. For example, if $u \subseteq t$ and $s \sqsubseteq pat(u)$, then $NE_M(u, t)$ iff $NE_M(s, t)$. Note however that if a component s is below a component o, then $UN_M(o, t)$ does not necessarily imply $UN_M(s, t)$, although inessentiality of the subterm corresponding to o implies that of the subterm corresponding to e. Take for example $R = \{f(x) \to g(x),\ a \to b\}$, and take for M the set of terms not containing occurrences of a. Then a is M-needed in $f(a)$, while $f(a)$ is not. But we still have the following crucial property, used in many proofs of normalization of needed strategy:

(β) if $t \xrightarrow{u} t'$, $UN_M(u, t)$, and $u' \subseteq t'$ is a new redex, then $UN_M(u', t')$.

This follows from the fact that if P suppresses $pat(u)$, then $u + P/u$ suppresses $pat(u')$.

In Kennaway and Sleep [3] a labelling is defined for OCRSs, based on the labelling system of Klop [7], which is in turn a generalization of the labelling system for the λ-calculus introduced by Lévy [9]. To each labelled redex u an *index* $Ind(u)$ is associated in such a way that:

(a) [3] If a step $t \xrightarrow{u} s$ creates v, then $Ind(u) < Ind(v)$, and if $w \subseteq s$ is a rseidual of a redex $w' \subseteq t$, then $Ind(w) = Ind(w')$.

(b) [7] Any reduction in which only redexes with a bounded redex-index are contracted is terminating.

Theorem 5.1 (**Uniform Normalization**) Let M be a stable set of terms in an OERS R.

(1) Any term t in R not in M-normal form contains at least one M-needed redex.

(2) A term t in R has an M-normal form iff every M-needed reduction starting from t is terminating.

Proof (1) If t does not have M-normal form, then (1) is clear. So let $P : t \twoheadrightarrow s \xrightarrow{u} e$ be an M-normalizing reduction with $s \notin M$. By the stability of M, u is M-needed. By (α) and (β), it is either created by or is a residual of an M-needed redex, and (1) follows by repeating the argument.

(2) Let $P : t \twoheadrightarrow s$ be an M-normalizing reduction and $Q : t \xrightarrow{u_0} t_1 \xrightarrow{u_1} \ldots$ be an M-needed reduction. Further, let $P_i = P/(u_0 + \ldots + u_{i-1})$. By ($a$), $Ind(P_i) \leq Ind(P)$ ($Ind(P)$ is the maximal index of redexes contracted in P). Since Q is M-needed and P_i is M-normalizing, at least one residual of u_i is contracted in P_i. Therefore, again by (a), $Ind(u_i) \leq Ind(P_i)$. Hence $Ind(Q) \leq Ind(P)$ and Q is terminating by (b).

One can show using (γ) (e.g., as in [3] or in [4]) that, for any such stable set M that M-unneeded redexes cannot duplicate or erase M-needed ones, the M-needed strategy is *M-hypernormalizing.* That is, a term is M-normalizing iff it does not posses a reduction not containing terms from M

in which infinitely many times M-needed redexes are contracted. However, this is not the case for all stable M. Indeed, consider the OTRS $R = \{f(x) \to h(f(x), f(x)), a \to b\}$ and take for M the set of terms not containing occurrences of a. Then $f(a) \to h(f(a), f(a)) \to h(f(b), f(a)) \to h(f(b), h(f(a), f(a))) \to h(f(b), h(f(b), f(a))) \to \ldots$ contracts infinitely many M-needed redexes, while $f(a) \to f(b)$ is M-normalizing.

References

[1] Barendregt H. P., Kennaway J. R., Klop J. W., Sleep M. R. Needed Reduction and spine strategies for the lambda calculus. Information and Computation, v. 75, no. 3, 1987, p. 191-231.

[2] Huet G., Lévy J.-J. Computations in Orthogonal Rewriting Systems. In: Computational Logic, Essays in Honour of Alan Robinson, ed. by J.-L. Lassez and G. Plotkin, MIT Press, 1991.

[3] Kennaway J. R., Sleep M. R. Neededness is hypernormalizing in regular combinatory reduction systems. Preprint, School of Information Systems, University of East Anglia, Norwich, 1989.

[4] Z. Khasidashvili. β-reductions and β-developments of λ-terms with the least number of steps. In Proc. of the International conference in Computer Logic “COLOG 88”, Tallinn 1988, Springer LNCS, v. 417, 1990, p. 105–111.

[5] Khasidashvili Z. The Church-Rosser theorem in Orthogonal Combinatory Reduction Systems. Report 1825, INRIA Rocquencourt, 1992.

[6] Khasidashvili Z. Optimal normalization in orthogonal term rewriting systems. In: Proc. of the fifth International Conference on Rewriting Techniques and Applications, Springer LNCS, vol. 690, C. Kirchner, ed. Montreal, 1993, p. 243-258.

[7] Klop J. W. Combinatory Reduction Systems. Mathematical Centre Tracts n. 127, CWI, Amsterdam, 1980.

[8] Klop J. W., van Oostrom V., van Raamsdonk F. Combinatory reduction Systems: introduction and survey. In: To Corrado Böhm, J. of Theoretical Computer Science 121, 1993, p. 279-308.

[9] Lévy J.-J. Réductions correctes et optimales dans le lambda-calcul, Thèse de l'Université de Paris VII, 1978.

[10] Maranget L. La stratégie paresseuse, Thèse de l'Université de Paris VII, 1992.

[11] Nöcker E. Efficient Functional Programming. Compilation and Programming Techniques. Ph. D. Thesis, Katholic University of Nijmegen, 1994.