

Combinatory Reduction Systems with Explicit Substitution

Kristoffer H. Rose
*kris@diku.dk**

Extended Abstract, September 1, 1995

Abstract

We show how the idea of explicit substitution can be applied to give a measure of the computational complexity of rewriting in combinatory reduction systems (CRS) by constructing an extension R_x with explicit substitution for any CRS R . We prove that R_x is confluent if R is orthogonal.

Key words: orthogonal combinatory reduction systems, explicit substitution, confluence.

1 Introduction

The λ -calculus is a useful basis for denotational models of the computable functions. However, the tradition of counting β -reduction steps $(\lambda x.M)N \rightarrow M[x := N]$, where $M[x := N]$ is metanotation for the λ -term obtained by substituting in M all (unbound) occurrences of x by N , does not account for the size of the abstraction body M . This means that it is not a useful operational model of computation because the number of β -reduction steps required to model a computation does not correspond to the ‘work’ a real computer has to do.

The traditional solution to this problem has been to assume a *fixed* reduction strategy, and then invent abstract machines realising it in an operationally realistic way (Landin 1964, Henderson 1980). This is unsatisfying because it makes reasoning about the relative operational behaviour of reduction strategies difficult, *e.g.*, for showing the correctness of program transformations. A better solution in this respect is the study of explicit (or stepwise) substitution (de Bruijn 1978, Abadi, Cardelli, Curien and Lévy 1991, Kamareddine and Nederpelt 1993, Lescanne 1994) where each reduction step corresponds to a ‘realistic’ computation step without sacrificing the freedom to choose the reduction strategy.

Quite independent of this, Klop (1980) developed *combinatory reduction systems* (CRS) to provide a syntactic model for more general reduction systems with binding than pure λ -calculus. It seems natural to make this computational in the same sense as the λ -calculus by making explicit the steps involved in substitution. There is a technical problem, though: the explicit substitution tradition quoted above is based on λ -calculus with indices à la de Bruijn (1972) but CRS uses ‘real’ variables in an essential way.

* DIKU, University of Copenhagen, Universitetsparken 1, DK-2100 København Ø; (URL: <http://www.diku.dk/~kris/>).

Overview of the paper. Hence in this paper we study how the *named explicit substitution* principle of Rose (1992) can be applied to CRS, by systematising the way this was done for the λ xgc-calculus of Rose and Bloo (1995), which we reproduce as a case study.

We assume familiarity with the λ -calculus (Barendregt 1984) as well as some knowledge of higher order rewriting (we will use the CRS formalism as described by Klop, van Oostrom and van Raamsdonk (1993)).

2 Combinatory Reduction Systems

The intention with CRS is to add ‘bound variables’ and the associated concept of ‘substitution’ to TRS, thus combining TRS with λ -calculus – in fact, the initial treatment of Klop (1980) (as well as the investigation of Aczel (1978)) was concerned with ‘definable extensions of the λ -calculus’.

This section gives an introduction to the notion of *Combinatory Reduction System* (CRS). We first provide some intuition and then proceed with the needed formal results. The presentation follows the (recommended) (Klop et al. 1993) closely, with a slight change of notation to facilitate induction over syntax.

We start with the complete definition followed by a discussion of some important issues, then we state the results we will be using.

2.1 Definition (Combinatory Reduction System). A *CRS* is defined by

- A set of *function symbols*, aka the *alphabet*, denoted F, G, \dots (and sometimes greek letters and weird symbols); each with a *fixed arity* added as superscript (*i.e.*, F^n).
- The corresponding *metaterms*, denoted t, s, \dots , have the form

$$t ::= x \mid [x]t \mid F^n(t_1, \dots, t_n) \mid Z^n(t_1, \dots, t_n)$$

called *variables*, *metaabstractions*, *constructions*, and *metaapplications*, respectively, where x, y, z, \dots are used to denote (simple) *variables* and Z, Y, X (only!) to denote *metavariables*, each with an explicit arity superscript as indicated (and possibly marks, subscripts, *etc.*). We extract the *set of metavariables* of a term with $mv(\cdot)$.

Furthermore, metaterms must always be *closed*: have no free variables, *i.e.*, $fv(t) = \emptyset$ where $fv(x) = \{x\}$, $fv([x].t) = fv(t) \setminus \{x\}$, $fv(F^n(t_1, \dots, t_n)) = fv(t_1) \cup \dots \cup fv(t_n)$, and $fv(Z^n(t_1, \dots, t_n)) = fv(t_1) \cup \dots \cup fv(t_n)$.

- Metaterms are always considered modulo renaming of bound variables: we take as implicit the usual α -equivalence \equiv (defined inductively by $x \equiv x$, $[x]t \equiv [y]s$ if $t[x := z] \equiv s[y := z]$ for $z \notin fv(st)$, and $F(s_1, \dots, s_n) \equiv F(t_1, \dots, t_n)$ if $s_i \equiv t_i \ \forall i$; the renaming $t[x := z]$ is defined the usual way as $x[y := z] = z$ if $x = y$, $x[y := z] = x$ if $x \neq y$, $([x]t)[y := z] = [x']t[x := x'] [y := z]$ with $x' \notin fv([x]t) \cup \{y, z\}$, and
- Rewrite rules*, written $p \rightarrow t$, with the following restrictions:
 - The LHS (left hand side) p must be a *linear pattern*: it should be a construction, all metavariables should be distinct, and each metaabstraction should have pairwise distinct variables as parameters.
 - All metaapplications on the RHS (right hand side) t must use metavariables that occurred in the LHS with the same arity.

2.2 Remark (Term Rewriting Systems). Any term rewrite system (TRS) is also a CRS containing no metaabstractions and only metavariables of arity zero (namely the TRS variables). The restriction to terms of the defined relation is the same as defined by the TRS, so in general we call any metaterm containing no metaabstractions for a *term*.

2.3 Example ($\lambda\beta$ -calculus). The ordinary untyped $\lambda\beta$ -calculus is described by the CRS with function symbols $\{\lambda^1, @^2\}$ and the single rule $@^2(\lambda^1([x]Z^1(x)), Y^0()) \rightarrow Z^1(Y^0())$. This corresponds to the usual definition by substitution in that the bound variable to be substituted, x , is represented implicitly on the right side through the use of the metavariable Z^1 – informally we could have written $Z^1(Y^0())$ with ordinary substitution as something like $(Z^1(x))[x := Y^0()]$.

2.4 Notation (CRS abbreviations). We will use the usual CRS abbreviations:

- The arity superscript of function symbols and metavariables is omitted when obvious, and we omit the $()$ after zero-arity symbols.
- $[x, y, z]t$ abbreviates $[x][y][z]t$.
- $Fxyz.t$ similarly abbreviates ‘F-abstraction’ $F([x][y][z]t)$.
- st abbreviates ‘application’ $@(s, t)$ (when it has no other interpretation) where $@^2$ is then included as a function symbol.
- \vec{x}_n, \vec{Z}_n , and \vec{t}_n , abbreviate the sequences $x_1, \dots, x_n, Z_1, \dots, Z_n$, and t_1, \dots, t_n , respectively.

We use $()$ to disambiguate where needed, and application binds closer than abstraction. This allows ordinary conventions to be followed: $\lambda xyz.xz(yz)$ denotes $\lambda^1([x][y][z]@^2(@^2(x, z), @^2(y, z)))$.

2.5 Example (readable λ -calculus). Using the above abbreviations we can express the λ -calculus in a more usual notation: the following CRS is exactly the same as the previous:

$$(\lambda x.Z(x))Y \rightarrow Z(Y) \quad (\beta)$$

(but we could not have written the RHS as ZY , of course).

2.6 Example (combinatory logic). The class of *applicative* TRS is naturally expressed using the above abbreviation. An example is combinatory logic:

$$SXYZ \rightarrow (XZ)(YZ) \quad (S)$$

$$KXY \rightarrow X \quad (K)$$

(but we should not write the RHS of (S) as $XZ(YZ)$ because that can be interpreted as $@(X, Z(@^2(Y, Z)))$ instead of the intended $@(@^2(X, Z), @^2(Y, Z))$.)

The substitution concept of CRS is reminiscent of two-level λ -calculus in that metavariables are always ‘metaapplied’ to the list of terms that should be substituted into it. Metavariables are therefore instantiated to ‘metaabstractions’ denoted $\underline{\lambda}\vec{x}_n.t$ and the resulting ‘metaredex’ play the rôle of substitution.

2.7 Definition (substitution). A *valuation* σ assigns to each metavariable Z a *substitute* $\Delta(x_1, \dots, x_n).t$ where the x_i are distinct. Valuations are homomorphically extended to metaterms: $\sigma(t)$ denotes the result of first inserting $\sigma(Z)$ for each metavariable Z in t and then replace all *substitutions* $(\lambda(\vec{x}_n).t)(\vec{t}_n)$ by $t[x_1 := t_1, \dots, x_n := t_n]$ defined inductively by

$$\begin{aligned} x_i[x_1 := t_1, \dots, x_n := t_n] &\equiv t_i \\ x[x_1 := t_1, \dots, x_n := t_n] &\equiv x \quad x \notin \{x_1, \dots, x_n\} \\ ([x]t)[x_1 := t_1, \dots, x_n := t_n] &\equiv [x](t[x_1 := t_1, \dots, x_n := t_n]) \quad x \notin \{x_1, \dots, x_n\} \\ F^m(s_1, \dots, s_m)[x_1 := t_1, \dots, x_n := t_n] &\equiv F^m(s_1[x_1 := t_1, \dots, x_n := t_n], \dots, s_m[x_1 := t_1, \dots, x_n := t_n]) \end{aligned}$$

2.8 Remark. A lot of complexity is hidden in the requirement that valuations be ‘homomorphically extended’ to metaterms because of the risk of name clashes. This is solved by requiring that each combination of a rule and valuation is *safe* by renaming.¹

2.9 Definition (reduction). Given a rule $s \rightarrow t$. A *redex* of this rule is a term $\sigma(s)$ for some valuation σ , $\sigma(s) \rightarrow \sigma(t)$ is a *rewrite*, and for any context $C[\]$, $C[\sigma(s)] \rightarrow C[\sigma(t)]$ is a *reduction*.

Our attention in this paper is restricted to the orthogonal CRSs because confluence is then secured.

- 2.10 Definitions (orthogonal CRS).**
- a. A CRS is *nonoverlapping* if for any redex r containing another, this other always appears as a subterm of something matched to a variable.
 - b. A CRS is *left-linear* if all LHSs contain at most one occurrence of each metavariable.
 - c. A CRS is *orthogonal* if it is left-linear and nonoverlapping.

2.11 Theorem (confluence of orthogonal CRS). *Any orthogonal CRS is confluent.*

3 Explicit Substitutes

In this section we show how the substitutes of CRSs can be made explicit.

First we demonstrate the idea as it was first developed for a particular CRS, namely the $\lambda\beta$ -calculus. Then we identify a subclass of CRS that have ‘explicit substitution’ and show that every (orthogonal) CRS can be transformed into a (confluent) CRS in this class, thus providing an operationally equivalent system for which the complexity of reduction can be investigated in a realistic way.

3.1 Example ($\lambda\beta$ -calculus). As mentioned above, the idea of explicit substitution stems from studies of the λ -calculus. We demonstrate this by showing λxgc of Rose and Bloo (1995): it uses variables and can thus be written as the CRS (with explicit application operator $@^2$)

$$\begin{aligned} @((\lambda x.Z(x)), Y) &\rightarrow \Sigma([x]Z(x), Y) & (b) \\ \Sigma([x]x, Y) &\rightarrow Y & (xv) \\ \Sigma([x]Z, Y) &\rightarrow Z & (xgc) \\ \Sigma([x]\lambda y.Z(x, y), Y) &\rightarrow \lambda[y](\Sigma([x]Z(x, y), Y)) & (xab) \\ \Sigma([x]@(Z_1(x), Z_2(x)), Y) &\rightarrow @(\Sigma([x]Z_1(x), Y), \Sigma([x]Z_2(x), Y)) & (xap) \end{aligned}$$

¹The main complexity in the author’s CRS interpreter is this renaming.

It is instructive to compare this system to the definition of λxgc (Rose and Bloo 1995): they are identical except for syntax conventions and an important point: the requirement that all terms are closed prohibits free variables in CRS rules which means that all free variables must be ‘hidden’ inside metavariables (to which they are not given as parameters). This means that we cannot distinguish between substituting a free variable (λxgc rule ($xvgc$)) and the more general garbage collection rule (xgc).

Notice that this is not an orthogonal CRS: it has an overlap between the first and last rule with terms of the form $\Sigma([x]@(\lambda[y]Z_1(x, y), Z_2(x)), Y)$. However, it is confluent as shown in Rose and Bloo (1995) (and below as a consequence of the general case).

The key observation in the development for the λ -calculus is that no knowledge of the ‘depth’ of terms is needed in order to reduce the rules. In the example rules above this is manifest in the fact that all metaabstractions on both LHS and RHS have *only variable arguments*, in fact the *same variables on both sides*!

Analogously to the λ -calculus, the operational problem of using CRS reduction steps as a complexity measure is that the complexity of substitution depends on the size of the body of the substitute. The solution is also the same: to perform substitution in a stepwise manner, or, put differently, to change metareduces to use *explicit substitutes*, such that only ‘superficial term knowledge’ is used in the rules, and it is obtained in the same manner, by ‘unfolding’ the definition of substitutes into the rule set.

3.2 Definition (explicit substitution CRS). A CRS R is an *explicit substitution CRS* or simply *XCRS*, if all metaapplications on the RHS match the form $[x_1, \dots, x_n]Z(x_1, \dots, x_n)$. Individual metaabstractions on this form in RHSs will be called *explicit metaabstractions*.

3.3 Remark. Notice that even if it does not quite look so, the above system is on this form. This comes from the fact of CRS that *only bound variables are mentioned*. Specifically, the underlined part of $\lambda[y](\Sigma([x]Z(x, y), Y))$, the RHS of (xab), is of the form $[x]Z(x)$, because the y variable is external to this part. And renaming does not matter, so it is also of the form $[y]Z(y)$, *etc.*

Clearly the λxgc -calculus presented above is an XCRS. The interesting observation which we can carry over from the study of that calculus is that substitution does not depend on the context in which it is created.

This insight can be used to create an XCRS automatically from a CRS in a manner that ensures that the properties are preserved. The rest of this section details that construction.

3.4 Definition (CRS explicification). Given a CRS R with alphabet F_1^n . The CRS R_x is obtained by the steps listed in figure 1. If R defined the relation \rightarrow then we will denote the relation defined by R_x as \xrightarrow{x} ; the subrelation consisting of only the introduction rules (with name ($r-x$)) is denoted \xrightarrow{xI} , and the subrelation containing only the distribution and elimination rules (with names ($x-*$)) is denoted \xrightarrow{xE} .

3.5 Example. The system $\lambda\beta x$ generated this way for the $\lambda\beta$ -calculus with the rule

$$(\lambda x.Z(x))Y \rightarrow Z(Y) \quad (\beta)$$

Substitution introduction. For each rule (r) of the CRS construct a new rule (r-x) by replacing in the RHS all non-explicit metaabstractions $Z^n(\vec{t}_n)$ by the (explicit) metaabstraction

$$\Sigma^{n+1}([\vec{x}_n]Z(\vec{x}_n), \vec{t}_n)$$

where Σ^{n+1} is a new function symbol added to the alphabet (thus there will be one new symbol for each metavariable arity).

Stepwise substitution distribution. For each possible pair of Σ^{n+1} and F^m , add the rule

$$\begin{aligned} \Sigma^{n+1}([\vec{x}_n]F^m(Z_1(\vec{x}_n), \dots, Z_m(\vec{x}_n)), \vec{Y}_n) \\ \rightarrow F^m(\Sigma^{n+1}([\vec{x}_n]Z_1(\vec{x}_n), \vec{Y}_n), \dots, (\Sigma^{n+1}([\vec{x}_n]Z_m(\vec{x}_n), \vec{Y}_n))) \end{aligned} \quad (\text{x-F-n})$$

(n rules are added for each function symbol F^m).

Substitution elimination. For each of these new Σ^{n+1} add the rules

$$\begin{aligned} \Sigma^{n+1}([\vec{x}_n]x_1, \vec{Z}_n) &\rightarrow Z_1 & (\text{xv-n-1}) \\ &\vdots & \vdots \\ \Sigma^{n+1}([\vec{x}_n]x_n, \vec{Z}_n) &\rightarrow Z_n & (\text{xv-n}) \\ \Sigma^{n+1}([\vec{x}_n]Y, \vec{Z}_n) &\rightarrow Y & (\text{xgc-n}) \\ \Sigma^{n+1}([\vec{x}_n]Y(\vec{x}_n, y), \vec{Z}_n) &\rightarrow [y]\Sigma^{n+1}([\vec{x}_n]Y(\vec{x}_n, y), \vec{Z}_n) & (\text{xma-n}) \end{aligned}$$

(a total of $n + 2$ rules for each Σ^{n+1} are added).

Figure 1: Explicification steps.

is the following:²

$$\begin{aligned} (\lambda x. Z(x))Y &\rightarrow \Sigma([x]Z(x), Y) & (\beta\text{-x}) \\ \Sigma([x](Z_1(x))(Z_2(x)), Y) &\rightarrow (\Sigma([x]Z_1(x), Y))(\Sigma([x]Z_2(x), Y)) & (\text{x-@-1}) \\ \Sigma([x]\lambda(Z(x)), Y) &\rightarrow \lambda(\Sigma([x]Z(x), Y)) & (\text{x-}\lambda\text{-1}) \\ \Sigma([x]x, Y) &\rightarrow Y & (\text{xv-1-1}) \\ \Sigma([x]Z, Y) &\rightarrow Z & (\text{xgc-1}) \\ \Sigma([x, y]Z(x, y), Y) &\rightarrow [y]\Sigma([x]Z(x, y), Y) & (\text{xma-1}) \end{aligned}$$

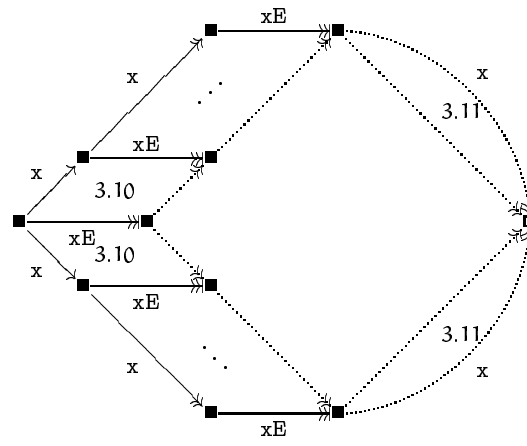
It is the same as the λxgc CRS shown above except that the abstraction distribution rule (xab) has been split into two steps: (x- λ -1) and (xma-1).

3.6 Remark. Clearly Rx is always an XCRS: the procedure is idempotent when applied to an XCRS because nothing is added when no non-explicit metaabstractions exist. Also notice that the resolution and distribution rules only depend on the alphabet.

The following general result is easy to show:

²In fact generated automatically from (β) by a program developed by the author.

Proof. The following diagram exploits \rightarrow_{λ} to show \rightarrow_{λ}^* :



(The technique is called the ‘interpretation method’ by Hardin and Lévy (1990).)

□

4 Conclusions

We have generalised the notion of explicit substitution to CRS as requested by Klop et al. (1993, sec.15(h)).

Further directions. Current work includes generalising other results obtained for λxgc to XCRS:

- Preservation of SN: show that $SN_{R_x} = SN_R$.
- Derive abstract machines for orthogonal CRS through their XCRS.
- Investigate if a variable-free (de Bruijn) notation is feasible.
- Investigate whether cyclic substitution (Rose 1992, Ariola and Klop 1994) can be integrated.
- Investigate whether the technique is applicable to other forms of higher order rewriting, in particular the relation to the ‘substitution calculus’ of van Oostrom and van Raamsdonk (1995).
- Also the relation to functional programming language, in particular supercombinator reduction, is under investigation.

References

- Abadi, M., Cardelli, L., Curien, P.-L. and Lévy, J.-J. (1991). Explicit substitutions. *Journal of Functional Programming* 1(4): 375–416.
- Aczel, P. (1978). A general Church-Rosser theorem. *Technical report*. Univ. of Manchester.
- Ariola, Z. M. and Klop, J. W. (1994). Cyclic lambda graph rewriting. *Proceedings, Ninth Annual IEEE Symposium on Logic in Computer Science*. Paris, France. pp. 416–425.
- Barendregt, H. P. (1984). *The Lambda Calculus: Its Syntax and Semantics*. Revised edn. North-Holland.

- de Bruijn, N. G. (1972). Lambda calculus notation with nameless dummies, a tool for automatic formula manipulation with application to the Church-Rosser theorem. *Koninklijke Nederlandse Akademie van Wetenschappen, Series A, Mathematical Sciences* 75: 381–392.
- de Bruijn, N. G. (1978). A namefree lambda calculus with facilities for internal definition of expressions and segments. *TH-Report 78-WSK-03*. Dept. of Mathematics, Technological University Eindhoven. Netherlands.
- Hardin, T. and Lévy, J.-J. (1990). A confluent calculus of substitutions. *Rapport de Recherche 90-11*. CEDRIC. 292, rue Saint-Martin, 75141 Paris CEDEX 03. Also in France-Japan Artificial Intelligence and Computer Science Symposium, Izu, 1989.
- Henderson, P. (1980). *Functional Programming—Application and Implementation*. Prentice-Hall.
- Kamareddine, F. and Nederpelt, R. (1993). On stepwise explicit substitution. *International Journal of Foundations of Computer Science* 4(3): 197–240.
- Klop, J. W. (1980). *Combinatory Reduction Systems*. Mathematical Centre Tracts 127. Mathematisch Centrum, Amsterdam.
- Klop, J. W., van Oostrom, V. and van Raamsdonk, F. (1993). Combinatory reduction systems: Introduction and survey. *Theoretical Computer Science* 121: 279–308.
- Landin, P. (1964). The mechanical evaluation of expressions. *Computer Journal* 6: 308–320.
- Lescanne, P. (1994). From $\lambda\sigma$ to $\lambda\nu$: a journey through calculi of explicit substitutions. *POPL '94—21st Annual ACM Symposium on Principles of Programming Languages*. Portland, Oregon. pp. 60–69.
- Rose, K. H. (1992). Explicit cyclic substitutions. In M. Rusinowitch and J.-L. Rémy (eds), *CTRS '92—3rd International Workshop on Conditional Term Rewriting Systems*. Number 656 in *LNCS*. Springer-Verlag. Pont-a-Mousson, France. pp. 36–50. [URL: ftp://ftp.diku.dk/diku/semantics/papers/D-143.ps](ftp://ftp.diku.dk/diku/semantics/papers/D-143.ps)
- Rose, K. H. and Bloo, R. (1995). Named lambda calculi with explicit substitution and garbage collection. *Semantics note*. DIKU (University of Copenhagen). In preparation; available through [URL: http://www.diku.dk/~kris/research.html](http://www.diku.dk/~kris/research.html).
- van Oostrom, V. and van Raamsdonk, F. (1995). Weak orthogonality implies confluence: the higher-order case. *Technical Report CS-R9501*. CWI.