Semantic Labelling for Termination of Combinatory Reduction Systems

Makoto Hamana
Department of Computer Science, Gunma University, Japan
hamana@cs.gunma-u.ac.jp

Abstract. We give a method of proving termination of Klop’s higher-order rewriting format, combinatory reduction system (CRS). Our method called higher-order semantic labelling is an extension of Zantema’s semantic labelling for first-order term rewriting systems. We systematically define the labelling by using the complete algebraic semantics of CRS.

1 Introduction

Rewrite rules appear everywhere in computer science and logic. When reasoning with such rewrite rules, termination property is important, so we need some way to ensure termination of rewrite rules. This topic has been extensively investigated in the field of term rewriting [3]. Higher-order extension of term rewriting are known as several formats: major representatives are Klop’s Combinatory Reduction Systems (CRSs) [13], Nipkow’s Higher-order Rewrite Systems [15], and Blanqui, Okada and Jouannaud’s systems [5, 4]. In this paper, we deal with Klop’s CRSs. We give a method to prove termination (meaning strong normalisation) of a CRS, called higher-order semantic labelling. This is an extension of semantic labelling for first-order term rewriting systems (TRSs) given by Zantema [17].

Background. This paper is based on the algebraic semantics of CRS [11] by $\Sigma$-monoids. The notion of $\Sigma$-monoids was introduced by Fiore, Plotkin and Turi [9], then a higher-order abstract syntax for free $\Sigma$-monoids was developed by the author [10]. Full version paper is available from http://www.keim.cs.gunma-u.ac.jp/hamana/hslf.pdf.

2 Preliminaries

CRS. We use the definition of the standard reference [14] of CRSs with a slight modification of syntax used in [8]. This is the exactly the same one used in [11]. Moreover, we treat only CRSs built from binding signatures, which we call structural CRSs [11]. Hereafter, we do not explicitly say “structural”. A binding signature $\Sigma$ is consisting of a set $\Sigma$ of function symbols with an arity function $a : \Sigma \to \mathbb{N}$ ($\mathbb{N}$ denotes the set of all finite sequences of natural numbers). A function symbol of binding arity $(n_1, \ldots, n_l)$, denoted by $f : (n_1, \ldots, n_l)$, has $l$ arguments and binds $n_i$ variables in the $i$-th argument ($1 \leq i \leq l$). For a formal treatment of named variables modulo $\alpha$-equivalence in CRSs, we assume the method of de Bruijn levels [9] for the naming convention of variables in CRSs. We also use the convention that $n \in \mathbb{N}$ denotes the set $\{1, \ldots, n\}$ ($n$ is possibly 0). Under the method of de Bruijn levels, this $n$ means the set of variables from 1 to $n$. 
Fix an \( \mathbb{N} \)-indexed set \( Z \) of metavariables defined by \( Z(l) = \{ z \mid z \text{ has arity } l \} \) (\( z \) is called a metavariable). A meta-term \( t \) is obtained if \( n \vdash t \) is derived from the following rules for some \( n \in \mathbb{N} \).

\[
\begin{align*}
x \in n & \quad f : \langle i_1, \ldots, i_\ell \rangle \in \Sigma \quad n + i_1 \vdash t_1 \cdots n + i_\ell \vdash t_\ell \quad z \in Z(l) & \quad n \vdash t_1 \cdots n \vdash t_\ell \\
 n + x & \quad n \vdash f(n+1 \ldots n+i_1, t_1, \ldots, n+1 \ldots n+i_\ell, t_\ell) & \quad n \vdash Z(t_1, \ldots, t_\ell)
\end{align*}
\]

A meta-term containing no metavariables is called a term. The \( \mathbb{N} \)-indexed set of all metaterms is defined by \( M_\Sigma Z(n) = \{ t \mid n \vdash t \} \) generated by an \( \mathbb{N} \)-indexed set of metavariables \( Z \). A rewrite rule, written \( l \rightarrow r \), consists of two meta-terms \( l \) and \( r \) with the following additional restrictions: (i) \( l \) and \( r \) are closed (w.r.t. variables) meta-terms, (ii) \( l \) must be a function term where all meta-applications have the form \( z[x_1, \ldots, x_n] \) with distinct variables \( x_i \), (iii) \( r \) can only contain meta-applications with meta-variables occurring in the left-hand side. We may use the notation \( Z\langle n \rangle \vdash s \rightarrow t \) for a rule or a rewrite step if metavariables and variables in \( s \) and \( t \) are included in \( Z \) and \( n \) respectively. We may also simply write \( Z \vdash s \rightarrow t \) or \( n \vdash s \rightarrow t \) if another part is not important. A valuation \( \theta \) is a mapping that assigns to \( n \)-ary metavariable \( z \) a term \( t \) \( \theta : z \mapsto A(x_1, \ldots, x_n) \) where all variables in \( t \) are included in \( \{ x_1, \ldots, x_n \} \). Any valuation is naturally extended to a function on meta-terms. A set of rewrite rules under the signature \( \Sigma \) is called a CRS and denoted by \((\Sigma, \mathcal{R})\) or simply \( \mathcal{R} \). The CRS rewrite relation \( \rightarrow_\mathcal{R} \) is generated by context and safe valuation closure of a given CRS \( \mathcal{R} \).

**Binding algebras.** Algebraic semantics of CRSs is given by binding algebras. Let \( \mathcal{F} \) be the category which has finite cardinals \( n = \{ 1, \ldots, n \} \) \((n \text{ is possibly } 0)\) as objects, and all functions between them as arrows \( m \rightarrow n \). This is the category of object variables by the method of de Bruijn levels (i.e. natural numbers) and their renamings. We use the functor category \( \text{Set}^d \). An object of \( \text{Set}^d \) may be called a presheaf. The functor \( \delta : \text{Set}^d \rightarrow \text{Set}^d \) for “index extension” is defined by \( \delta L(n) = L(n+1) \) for \( L \in \text{Set}^d \).

To a binding signature \( \Sigma \), we associate the signature functor \( \Sigma : \text{Set}^d \rightarrow \text{Set}^d \) given by \( \Sigma A = \prod_{f : (n_1, \ldots, n_k) \in \Sigma} \prod_{i \in \Sigma} \delta^n A \). An \( \Sigma \)-algebra is an algebra of this functor. The presheaf of values \( V \in \text{Set}^d \) is given by \( V(n) = n \). The \( \mathbb{N} \)-indexed set of all metaterms \( M_\Sigma Z \) generated by \( Z \) forms a free \( \Sigma \)-monoid \([10]\). Moreover, when \( Z = 0 \) (empty set), \( M_\Sigma 0 \) forms an initial \( V + \Sigma \)-algebra and an initial \( \Sigma \)-monoid \([9, 10]\).

**Algebraic semantics of rewriting.** For a presheaf \( A \), we write \( \geq_A \) for a family of preorders \( (\geq_{A(n)})_{n \in \mathbb{N}} \), where \( \geq_{A(n)} \) is a preorder on set \( A(n) \) for each \( n \in \mathbb{N} \). Let \( (A_1, \geq_{A_1}), \ldots, (A_l, \geq_{A_l}), (B, \geq_B) \) be presheaves equipped with preorders. A map \( f : A_1 \times \cdots \times A_\ell \rightarrow B \) in \( \text{Set}^d \) is weakly monotone if all \( n \in \mathbb{N} \), all \( a_1, b_1 \in A_1(n), \ldots, a_\ell, b_\ell \in A_\ell(n) \) with \( a_k \geq_{A_k} b_k \) for some \( k \) and \( a_j = b_j \) for all \( j \neq k \), then \( f(n)(a_1, \ldots, a_\ell) \geq_B f(n)(b_1, \ldots, b_\ell) \). A weakly monotone \( V + \Sigma \)-algebra \((A, \geq_A)\) is a \( V + \Sigma \)-algebra equipped with preorders such that every operation is weakly monotone. For a \( V + \Sigma \)-algebra \( A \), a term-generated assignment \( \phi : Z \rightarrow A \) is a morphism of \( \text{Set}^d \) that is a composite \( Z \rightarrow M_\Sigma 0 \rightarrow A \) when the second morphism is a homomorphism. A \( V + \Sigma \)-algebra \( A \) satisfies a rewrite rule \( Z \vdash \overline{n} \cdot l \rightarrow \overline{n} \cdot r \) if \( \phi^n(n)(l) = \phi^n(n)(r) \) for all term-generated assignments \( \phi : Z \rightarrow A \). A model \( A \) for a CRS \((\Sigma, \mathcal{R})\) is a \( V + \Sigma \)-algebra \( A \) that satisfies all rules in the weakening closure \( \mathcal{R}^c \) (cf. \([11]\)) which means rules allow free variables in their instances. A weakly monotone \( V + \Sigma \)-algebra \((A, \geq_A)\) satisfies a rewrite rule \( Z \vdash \overline{n} \cdot l \rightarrow \overline{n} \cdot r \) if \( \phi^n(n)(l) \geq_A \phi^n(n)(r) \) for all term-generated assignments
\( \phi : Z \rightarrow A \). A quasi-model \( A \) for \((\Sigma, R)\) is a weakly monotone \( \Sigma + \Sigma\)-algebra \( A \) that satisfies all rules in the weakening closure \( R^\circ \).

3 Higher-Order Semantic Labelling

We assume that \( Z \) is an \( \mathbb{N} \)-indexed set of metavariables, \( \Sigma \) is a binding signature and \( M \) is a \( \Sigma + \Sigma \)-algebra. We introduce labelling of functions symbols: choose for every \( f \in \Sigma \) a corresponding non-empty set \( S_f \) of labels, called sort set. The binding signature \( \Sigma \) for labelled function symbols is defined by \( \Sigma = \{ f_p \mid f \in \Sigma, \ p \in S_f \} \) where the binding arity of \( f_p \) is defined to be the binding arity of \( f \). A function symbol is labelled if \( S_f \) contains more than one element. For unlabelled \( f \), the set \( S_f \) containing only one element can be left implicit; in that case we write \( f \) instead of \( f_p \). Choose for \( f' : (i_1, \ldots, i_l) \in \Sigma \), a sort map that is a morphism of \( \text{Set}^\Sigma \) defined by \( \delta^n M \times \cdots \times \delta^l M \rightarrow K_{S_f} \), where \( K_{S_f} \in \text{Set}^\Sigma \) is the constant presheaf defined by \( K_{S_f}(n) = S_f \). The sort map was originally called a projection, denoted by \( \pi_f \) in [17]. Then, as in the case of ordinary signature, we define \( M_{\Sigma} Z \) by the presheaf of all meta-terms generated by the labelled signature \( \Sigma \). Let \( \phi : Z \rightarrow M \) be an assignment. The labelling map \( \phi^l : M_{\Sigma} Z \rightarrow M_{\Sigma} Z \) is a morphism of \( \text{Set}^\Sigma \) defined by \( \phi^l_n(x) = x, \ \phi^l_n([l]) = z[\phi^l_n]^l \), \( \phi^l_n(f(n+1, \ldots, n+l, i_1, i_2)) = f(\phi^l_n(s_1), \ldots, \phi^l_n(s_l)) \). For a given CRS \((\Sigma, R)\), we define the labelled rules by \( \overline{R} = \{ Z + \overline{\pi}_f \phi^l_n \rightarrow \overline{\pi}_f \phi^l_n r \mid Z + \overline{\pi} l \rightarrow \overline{\pi} r \in R \} \), assignment \( \phi : Z \rightarrow M \). Thus \( \overline{R} \) is a set of rewrite rules on labelled terms in \( M_{\Sigma} Z(0) \). So, \( (\Sigma, \overline{R}) \) forms a CRS that gives rewriting on \( \Sigma \)-terms. The labelling map \( \phi^l \) preserves \( R \)-rewrite structures, hence we have the main theorem of this paper.

**Proposition 1.** Let \( M \) be a model of \((\Sigma, R)\). If we have CRS rewriting \( n + s \rightarrow_R t \) on \( M_{\Sigma} Z_0 \), then for the assignment \( \phi : 0 \rightarrow M \), we have rewriting \( n + \phi^k_n s \rightarrow_R \phi^k_n t \) on \( M_{\Sigma} Z_0 \).

**Theorem 2 (Higher-order semantic labelling).** Let \( M \) be a model of a CRS \((\Sigma, R)\). Then, \((\Sigma, R)\) is terminating if and only if \((\Sigma, \overline{R})\) is terminating.

We need separately a way to prove termination of the labelled system. For this purpose, we use Blanqui’s version of General Schema for “new definition of IDTS” ([6] Def. 1). Structural CRSs used in the present paper can be seen as a subclass of Blanqui’s IDTSs. Hence we can use General Schema as a criteria of termination of structural CRSs. In our experience, this is the most powerful decidable method to prove termination of CRSs. General Schema uses a precedence which is a partial order on function symbols occurring in a CRS.

**Example 3 (CRS for prefix sum).** Consider the following CRS \( P \) for computing the prefix sum of a list i.e. the list with the sum of all prefixes of a given list using the higher-order function map [7].

\[
\text{map}(a.R[a], \text{nil}) \rightarrow \text{nil} \quad \text{ps}(\text{nil}) \rightarrow \text{nil} \\
\text{map}(a.R[a], x : \text{xs}) \rightarrow t[x] : \text{map}(a.R[a], \text{xs}) \quad \text{ps}(x : \text{xs}) \rightarrow x : \text{ps}(\text{map}(a.R + a, \text{xs}))
\]
Unfortunately, General Schema cannot show termination of the CRS $\mathcal{P}$ because the argument of $\text{ps}$ in the right-hand side of the last rule is not a subterm of the argument of $\text{ps}$ in the left-hand side. So we use higher-order semantic labelling. The CRS $\mathcal{P}$ is formulated under the binding signature $\Sigma = \{\text{map} : \langle 1, 0 \rangle, \text{S}, \text{ps} : \langle 0 \rangle, 0, \text{nil} : \langle \rangle, +, \cdot : \cdot \rangle$. We take the presheaf $M_i : \{\text{map} \to N^i \to N\}$ of all functions on $\mathbb{N}$ for a model of $\mathcal{P}$.

The operations are defined by $\map_{M_0}(f, y) = y$, $\text{ps}(x) = x$, $\text{M}_0(x, y) = y + 1$. $\text{nil}_{\text{M}_0} = 0$, $x + \text{M}_0 y = 0$. The idea of this model is "to count the number of cons’s". We label the function symbol $\text{ps}$ and assume that other function symbols are unlabelled. We use the natural numbers $\mathbb{N}$ as the sort set $S_\text{ps}$. The sort map is defined by $\langle x \rangle_{\text{ps}}^0 = x$. Then, we have the labelled rules $\text{ps}_0(\text{nil}) \to \text{nil}$, $\text{ps}_{i+1}(x : \text{xs}) \to x : \text{ps}_i(\text{map}(a.x + a, \text{xs}))$ for all $i \in \mathbb{N}$. General Schema succeeds in showing termination of this labelled CRS with the precedence $\text{ps}_i > \text{ps}_{j} > \text{map} > \text{nil}$; for $i > j \in \mathbb{N}$.

The quasi-model version of the semantic labelling theorem is obtained similarly. Define the CRS $\text{Decr}$ (called “decreasing rules”) over $\overline{\Sigma}$ by $f_p(i_1^1.z_1^1[i_1^1], \ldots, i_l^1.z_l^1[i_l^1]) \to f_p(i_1^1.z_1^1[i_1^1], \ldots, i_l^1.z_l^1[i_l^1])$ for all $f : \langle i_1, \ldots, i_l \rangle \in \Sigma$ and all $p >_S q \in S_f$, where $>_S$ denotes the strict part of $\geq_S$.

**Theorem 4.** Let $M$ be a quasi-model for a CRS $(\Sigma, \mathcal{R})$ and $(\overline{\Sigma}, \overline{\mathcal{R}})$ the labelled CRS with respect to $M$. Then $(\Sigma, \mathcal{R})$ is terminating if and only if $(\overline{\Sigma}, \overline{\mathcal{R}} \cup \text{Decr})$ is terminating.

**Example 5 (Monad for the lambda calculus).** Consider the following CRS $\mathcal{R}$ under the binding signature $\Sigma = \{\text{lift}, \text{bind} : \langle 1, 0 \rangle, \text{app} : \langle 0, 0 \rangle, \text{abs} : \langle 1 \rangle, \text{var}, \text{old} : \langle 0 \rangle, \text{new} : \langle \rangle\}$ and the metavariables $Z = \{\! v \!, \! a \!, \! x \!, \! r \!, \! n \}$. This is the single base type version of the rewrite system of multiplication operation of the monad of $\lambda$-calculus in [1].

- $\text{lift}(\text{new}) \to \text{var}(\text{new})$
- $\text{lift}(\text{f}, \text{old}(\text{x})) \to \text{bind}(\text{a}, \text{var}(\text{old}(\text{a})), \text{r}(\text{x}))$
- $\text{bind}(\text{r}, \text{var}(\text{x})) \to \text{r}(\text{x})$
- $\text{bind}(\text{f}, \text{app}(\text{s}, \text{t})) \to \text{app}(\text{bind}(\text{f}, \text{s}), \text{bind}(\text{f}, \text{t}))$
- $\text{bind}(\text{r}, \text{abs}(\text{a}, \text{g}(\text{a}))) \to \text{abs}(\text{a}, \text{bind}(\text{b}, \text{lift}(\text{f}, \text{b}), \text{g}(\text{a})))$

It is not straightforward to show termination of this system using known criteria in higher-order rewriting. In the last rule, $\text{bind}$ decomposes the $\text{abs}$ construct, but in its right-hand side, the recursive call of $\text{bind}$ happens with the argument $h.\text{lift}(\text{f}, \text{b})$ which is structurally bigger than $\text{f}$ in the left-hand side. Moreover, $\text{lift}$ is defined by using $\text{bind}$, which seems to be circularity. Due to these reasons, most of RPO-like syntactical methods for termination of higher-order rewriting fail. When we try to show termination of $\mathcal{R}$ using General Schema, we cannot determine the precedence between the function symbols $\text{lift}$ and $\text{bind}$. This is the same as for the corresponding rewrite rules written in other formats of higher-order rewriting and termination criteria for them. For the higher-order rewrite system corresponding to $\mathcal{R}$ in the format Inductive Data Type Systems [5], the higher-order RPO [12] fails to show termination. For the $\text{S}$-expression rewrite system corresponding to $\mathcal{R}$, the lexicographic path order described in [16] fails to show termination. For the simply-typed term rewriting system (STTRS) corresponding to $\mathcal{R}$, the dependency pair technique for STTRSs [2] fails to show termination.

**Termination proof.** We take the carrier $M$ to be the presheaf of strictly monotone functions on $\mathbb{N}$ equipped with the usual order $\geq$ and its pointwise extension, i.e., $M_i = [\mathbb{N}^i \to \mathbb{N}]$ where $[- \Rightarrow -]$ denotes the set of all strictly monotone functions. This $M$ also forms...
a monoid in $\text{Set}$ by taking the multiplication as the composition and the unit as the projections. The $\Sigma$-algebra structure on $M$ is defined by $\text{lift}_{M}(f, x) = f(0)$, $\text{bind}_{M}(f, x) = f(x)+1$, $\text{new}_{M} = 0$, $\text{old}_{M}(x) = x$, $\text{app}_{M}(x, y) = \max(x, y)$, $\text{abs}_{M}(g) = 1$. One can easily check that $M$ with these operations is indeed a quasi-model for $\mathcal{R}$. We only label the function symbol $\text{bind}$. We choose the sort set $\Sigma_{\text{bind}} = \mathbb{N}$ and the weakly monotone sort map as $\langle - , - \rangle_{\text{bind}}^{0} : [\mathbb{N} \Rightarrow \mathbb{N}] \times \mathbb{N} \rightarrow \mathbb{N}$, $\langle f, x \rangle_{\text{bind}}^{0} = f(x)$. Using this, we have the following labelled CRS $\mathcal{R} \cup \text{Decr}$ for all strictly monotone functions $f \in [\mathbb{N} \rightarrow \mathbb{N}]$.

$$
\begin{align*}
\text{lift}(r, \text{old}(x)) &\rightarrow \text{bind}_{0}(a, \text{var}(\text{old}(a)), r[x]) \\
\text{bind}_{f_{(i, s, t)}}(f) &\rightarrow \text{app}(\text{bind}_{f_{(i, s, t)}}(f, s), \text{bind}_{f_{(i, s, t)}}(f, t)) \\
\text{bind}_{f_{(i, s, t)}}(f) &\rightarrow \text{app}(\text{bind}_{f_{(i, s, t)}}(f, s), \text{bind}_{f_{(i, s, t)}}(f, t)) \\
\text{bind}_{f_{(i, s, t)}}(f) &\rightarrow \text{abs}(a, \text{bind}_{f_{(i, s, t)}}(b, \text{lift}(f, b), g[a])) \\
\text{bind}(r, x) &\rightarrow \text{bind}_{(f, x)}(a) \\
\text{for all } i > j > 0.
\end{align*}
$$

With the precedence $\text{bind}_{i} > \text{bind}_{j} > \text{lift} > \text{bind}_{0} > \text{var}, \text{app}, \text{abs}, \text{old}, \text{new}$ for all $i > j > 0$. General Schema succeeds in showing the termination of $\mathcal{R} \cup \text{Decr}$. Hence, by Thm. 2, we conclude termination of $\mathcal{R}$.

References


