数理解析研究所講究録
第 625 巻 1987 年 233-241

Approximation Reduction and Approximation Rules of Term Rewriting Systems

直井徹　　　　稲垣康善

Tohru NAOI and Yasuyoshi INAGAKI
Faculty of Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, JAPAN

1. Introduction.

This paper formalizes the fixedpoint semantics of non-ambiguous linear term rewriting systems, which generalizes that of recursive program schemes. This is done by considering the infinite sequence of rewriting systems that approximate a given system. It also shows that the fixedpoint semantics coincides with the algebraic semantics previously proposed by the authors. This result generalizes the well-known fact that these two semantics for recursive schemes coincide [2].

Finally it gives some sufficient condition for the termination of the approximation term rewriting systems.

2. Preliminaries.

Infinite trees

Let F be a set of function symbols which are associated with some arities, and X be a set of variable symbols such that $F \cap X = \phi$. We assume F contains a special symbol Ω with arity 0. Let $T^{\infty}(F,X)$ be the set of finite or infinite trees which are well-formed with respect to the arity of symbols (in $F \cup X$) labeled at their nodes. An order $\leq$ on $T^{\infty}(F,X)$ is defined as follows: For T, $T' \in T^{\infty}(F,X)$, $T \leq T'$ iff T is obtained by substituting Ω's for some occurrences of subtrees in T'. It is known that $< T^{\infty}(F,X), \leq >$ is a cpo and the least element is Ω. See [1] for more details. The set of finite trees (terms) is denoted by T(F,X). We shall use t, t', u,..., for the elements of T(F,X) and T, T', U, ..., for the elements of $T^{\infty}(F,X)$. For a set of trees Δ, Δ^{-} denotes the closure of Δ defined as follows: let $\underline{\Delta} = \{t \mid \exists T \in \Delta \ t \leq T\}$ and $\Delta^{-} = \{\sqcup \Delta' \mid \Delta' \subseteq \underline{\Delta}$ directed$\}$.

Let P^{*} be the set of finite sequences of positive integers. The nodes of a tree are identified by elements of P^{*} in a well-known manner [1,4]. Hence, we can define the set of nodes in T, denoted by Dom(T), as a subset of P^{*}. For $p \in Dom(T)$, a tree T/p is the subtree of T whose root is the node p in T, and $T[p \leftarrow T']$ is the tree which is obtained by replacing the subtree of T occurred at p with T'.

Substitution is a mapping σ from X to $T^{\infty}(F,X)$. It is extended to a continuous mapping on $T^{\infty}(F,X)$ by $\sigma(T)=T[p \leftarrow \sigma(x) \mid T/p=x \in X]$.

Term Rewriting Systems

Let Var(T) be the set of variable symbols which occurs in T. A right-infinite term rewriting system (riTRS), is a subset R of $T(F,X) \times T^{\infty}(F,X)$ such that each $\langle t, T\rangle \in R$ satisfies $Var(T) \supseteq Var(t)$. An element $\langle t, T\rangle$ of R is called a rewrite rule. R is said to be a term rewriting system (TRS) if the right-hand side of each rule is also finite.

A reduction in R is a 4-tuple $\langle T, T', p, \langle u, U\rangle\rangle$ such that $p \in Dom(T)$, $\langle u, U\rangle \in R$, $T/p = \sigma(u)$ and $T' = T[p \leftarrow \sigma(U)]$ for some substitution σ. $\rightarrow_R$ is a binary relation on $T^{\infty}(F,X)$ obtained by dropping the third and fourth components of reductions.

A parallel reduction in R is a 4-tuple $\langle T, T', P, \rho\rangle$ where $P = \{p_1, p_2, \ldots,\}$ be a (possibly infinite) mutually disjoint subset of Dom(T) and $\rho = \{\langle t_1, T_1\rangle, \langle t_2, T_2\rangle, \ldots,\} \subseteq R$ such that there exist $\sigma_1, \sigma_2, \ldots,$ satisfying $T/p_i = \sigma_i(T_i)$ for all i, and $T' = T[p_i \leftarrow \sigma_i(T_i') \mid i=1,2,\ldots]$ [4]. The corresponding binary relation on $T^{\infty}(F,X)$ is denoted by $\twoheadrightarrow_R$.

We denote by ρ^* the reflexive-transitive closure of a binary relation ρ. A binary relation ρ on $T^{\infty}(F,X)$ is confluent iff the following condition holds: For every T, T_1, T_2, if T ρ^* T_1 and T ρ^* T_2, there exists T' such that T_1 ρ^* T' and T_2 ρ^* T'.

A tree T is linear if any variable symbol occurs in T at most once. A riTRS R is linear if for any $\langle t, T\rangle \in R$, t is linear.

T is Ω-free if there is no occurrence of Ω in T, and R is Ω-free if for any $\langle t, T\rangle \in R$, t is Ω-free.

T and T' is unifiable if $\sigma(T) = \sigma'(T')$ holds for some σ and σ'. R is said to be non-ambiguous (non-overlapping) if for every $\langle t, T\rangle, \langle t', T'\rangle \in$

R and every $p \in Dom(t)$, t/p and t' is unifiable iff $t/p \in X$.

In this paper, we assume riTRSs are linear, non-ambiguous and Ω-free. The following is an infinite-tree-rewriting version of the well-known result [4]:

Proposition 2.1. For every linear non-ambiguous riTRS R, $\twoheadrightarrow_R$ is confluent. And, for every linear non-ambiguous TRS R, $\rightarrow_R$ is confluent. □

3. *Algebraic Semantics*.

We briefly review the algebraic semantics of riTRSs. See [5] for the details.

We define the *set of redexes of* R, denoted by Red_R, as follows:

$$Red_R=\{T \mid T=\sigma(t) \text{ for some } \sigma \text{ and some } \langle u, V\rangle \in R\}.$$

The *set of candidates for redexes of* R, denoted by $Cand_R$, is defined inductively:

1) $T \in Red_R$ implies $T \in Cand_R$,

2) $T, T' \in Cand_R$ and $p \in Dom(T)$ implies $T[p \leftarrow T'] \in Cand_R$.

The *set of the candidates occurrences of* R *in* T, denoted by $Candocc_R(T)$, is defined by:

$$Candocc_R(T) = \{p \in Dom(T) \mid T/p \in Cand_R^{-}\}.$$

And the *set of approximation normal forms of* R, denoted by ANF_R, is defined by:

$$ANF_R = \{T \mid p \in Candocc_R(T) \text{ implies } T/p=\Omega\}.$$

A function ω_R on $T^{\infty}(F,X)$ is defined by

$$\omega_R(T)=T[p\leftarrow\Omega \mid p \text{ is outermost in } Candocc_R(T)].$$

$\omega_R(T)$ is called the approximation normal form of T w.r.t. R.

Definition 3.1. The algebraic semantics or the valuation of T by R is defined by:

$$Val_R(T) = \sqcup\{\omega_R(T') \mid T \twoheadrightarrow_R^* T'\}. \qquad \Box$$

Val_R is a retraction, i.e., continuous and idempotent, and its range is ANF_R. Using the continuity of Val_R, it is proved that $Val_R(T)=\sqcup\{\omega_R(T') \mid T \rightarrow_R^* T'\}$ holds for a TRS R (in [5], we adopted this as the definition of the definition of Val_R).

4. Fixedpoint Semantics.

We can take R as an equation system rather than a rewriting system and the left-hand sides as its unknown variables over ANF_R in the following manner. In sequel, we assume that R is a TRS. The set of symbolic interpretation of left-hand sides of R, denoted by Int_R, is the set of functions Θ from $\{t \mid \langle t, t'\rangle\in R\}$ to ANF_R such that $Var(t)\supseteq Var(\Theta(t))$ for any $\langle t, t'\rangle\in R$. For $\Theta\in Int_R$ and $\langle t, t'\rangle\in R$, $\Theta(t)$ is called the symbolic interpretation of t w.r.t. Θ. We want to extend the domain of Θ into $T^{\infty}(F,X)$ to interpret right-hand side of R. For the purpose, we regard Θ as a non-ambiguous linear and Ω-free riTRS. Then, we can show that $Val_\Theta(t)=\Theta(t)$ for $\langle t, t'\rangle\in R$. An interpretation Θ is called a symbolic solution of R if $\Theta(t)=Val_\Theta(t')$ for every $\langle t, t'\rangle\in R$.

We can solve the equation R by the fixedpoint theorem. First, we define a function Eq_R on Int_R by

$$Eq_R(\Theta) = \Theta' \text{ where } \Theta'(t) = Val_{\Theta}(t') \text{ for } \langle t, t'\rangle \in R.$$

Obviously, $\Theta \in Int_R$ is a <u>symbolic solution of</u> R iff it is a fixedpoint of Eq_R. We define an order $\leq$ on Int_R as follows: for $\Theta, \Theta' \in Int_R$, $\Theta \leq \Theta'$ if $\Theta(t) \leq \Theta'(t)$ for $\langle t, t'\rangle \in R$. It is shown that $\langle Int_R, \leq\rangle$ is a cpo with the least element $\perp$ such that $\perp(t) = \Omega$ for any $\langle t, t'\rangle \in R$. Moreover, the continuity of Eq_R can also be shown. Hence, by the fixedpoint theorem, its least fixed point Λ exists and is given by:

$$\Lambda = \sqcup\{Eq_R{}^n(\perp) \mid n \geqq 0\}.$$

<u>Definition</u> 4.1. The <u>fixedpoint semantics of a term</u> T <u>w.r.t.</u> R is defined by

$$Fix_R(T) = Val_{\Lambda}(T).$$ □

Let $Rn = Eq_R{}^n(\perp)$ for each n. Rn is called the <u>n-th standard approximation of</u> R. By the continuity of Eq_R, we have:

$$Fix_R(T) = \sqcup\{Val_{Rn}(T) \mid n \geqq 0\}.$$

Remark that, to calculate $Fix_R(T)$, we do not use R itself to rewrite terms but Rn's. It is easy to see that,

$$R0 = \{\langle t, \Omega\rangle \mid \langle t, t'\rangle \in R\} \text{ and}$$

$$Rn{+}1 = \{\langle t, Val_{Rn}(t')\rangle \mid \langle t, t'\rangle \in R\} \text{ for } n \geqq 0.$$

Intuitively, we at first interpret the left-hand side t as "undefined" and then, our approximate interpretation is refined asymptotically.

The fixedpoint semantics coincides with the algebraic semantics:

<u>Theorem</u> 4.2. $Fix_R = Val_R$. □

5. Termination on approximation normal forms

In this chapter, we study the termination of the approximation TRSs. Since we consider TRSs are abstract interpreters of programs, their termination can not always be assumed. Such an assumption put a severe restriction on our discussion. However, the termination of the approximation TRSs is not so restrictive as we see below.

Let Δ be a set of finite terms. We say that R is terminating over Δ if for any t in Δ there is no infinite reduction sequence issued from t. We also say R is terminating if R is terminating over T(F,X). A set of finite terms ANF_R^{rc} is defined inductively:

1) $ANF_R \cap T(F,X) \subseteq ANF_R^{rc}$,
2) t, $t' \in ANF_R^{rc}$ and $p \in Dom(t)$ implies $t[p \leftarrow t'] \in ANF_R^{rc}$.

Now we can state a sufficient condition for the termination of the approximation systems.

Proposition 5.1. If a TRS R is terminating over ANF_R^{rc}, its standard approximations are terminating TRSs. □

There is a sufficient condition for the termination of R over ANF_R^{rc}:

Proposition 5.2. Suppose R_1 is a terminating TRS and R_2 is a TRS such that a rule in R_2 has the form $\langle f(x_1,..,x_n), t\rangle$ and f does not occur in rules in R_1. Then a TRS $R = R_1 \cup R_2$ is terminating over ANF_R^{rc}. □

TRSs described in the above proposition still generalizes recursive program schemes.

6. Conclusion.

Through the limit of an infinite sequence of approximation rules, the fixedpoint semantics of term rewriting systems has been formalized as a generalization of that of recursive program schemes. On the other hand, the algebraic semantics is defined through the limits of infinite approximation reductions. However, as we have seen, these semantics coincide as like for recursive program schemes.

We also studied the termination of approximation TRSs. Terminating approximation systems will be a useful tool to check properties of the original system. Hence, a TRS which satisfies the condition of Proposition 5.2 can be said canonical in a sense.

Acknowledgment

The authors wish to express their gratitude to Dr. Namio HONDA, President of Toyohashi University of Technology and Dr. Teruo FUKUMURA, Professor of Nagoya University for their encouragements to conduct this work. They also thank Associate Professor Hirotomo ASO of Tohoku University and Associate Professor Toshiki SAKABE of Mie University for their helpful discussions. Their work is partly supported by the Grants from Ministry of

Education, Science and Culture of Japan (General(C) No. 60550263 and Special Project Research No. 61102003).

References

[1] B.Courcelle, Fundamental Properties of Infinite Trees, Theor. Comput. Sci. 25, pp. 95-169(1983).

[2] B.Courcelle and M.Nivat, Algebraic Semantics of Recursive Program Schemes, MFCS, Ed. Winkowski, Springer, LNCS 64, pp.16-30(1978)

[3] I.Guessarian and M.Nivat, About Ordered Sets in Algebraic Semantics, in: Orders: Description and Roles, Eds. M. Pouzet and D. Richard, North-Holland(1984).

[4] G.Huet, Confluent Reductions: Abstract Properties and Applications to Term Rewriting Systems, J.ACM 27, pp.797-821(1980).

[5] T.Naoi and Y.Inagaki, Algebraic Semantics of Term Rewriting Systems, Report of Tech. Group on Computation, COMP86-1, IECEJ(1986); Tech. Res. Rept. No.8603, Dept. of Information Science, School of Eng., Nagoya University(1986).

[6] T.Naoi and Y.Inagaki, The Relation between Algebraic and Fixedpoint Semantics of Term Rewriting Systems, Report of Tech. Group on Computation, COMP86-37, IECEJ(1986); Tech. Res. Rept. No.8604, Dept. of Information Science, School of Eng., Nagoya University(1986).