

# TYPE INFERENCE IN PROLOG AND ITS APPLICATION

Tadashi KANAMORI, Kenji HORIUCHI

Central Research Laboratory  
Mitsubishi Electric Corporation  
Tsukaguchi-Honmachi 8-I, Amagasaki, Hyogo, Japan 661

## Abstract

In this paper we present a type inference method for Prolog programs. The new idea is to describe a superset of the success set by associating a type substitution (an assignment of sets of ground terms to variables) with each head of definite clause. This approach not only conforms to the style of definition inherent to Prolog but also gives some accuracy to the types inferred. We show the basic computation method of the superset by sequential approximation as well as an incremental method to utilize already obtained results. We also show its application to verification of Prolog programs.

## 1. Introduction

It gives usefull information not only to programmers but also to met a-processing systems to infer characteristics of execution-time behavior from program texts. Especially in Prolog, it is usefull to know of what type a variable in a query is when the query succeeds. We call such a task *type inference* [8]. In this paper we present a type inference method for Prolog programs. After summarising preliminary materials in section 2, we describe type inference algorithms for Prolog programs in section 3. An application to verification is shown in section 4. See [7] for more details.

## 2. Preliminaries

In the followings, we assume familiarity with the basic terminologies of first order logic such as term, atom, substitution and most general unifier(mgu). We also assume knowledge about semantics of Prolog such as Herbrand universe  $H$ , Her brand interpretation  $I$ , minimum Herbrand Model  $M_0$  and transformation  $T$  of Herbrand interpretation associated with Prolog programs (see [11,14]). We follow the syntax of DEC-10 Prolog [10]. An atom  $p(X_1, X_2, \dots, X_n)$  is said to be *in general form* when  $X_1, X_2, \dots, X_n$  are distinct variables. A substitution  $\sigma$  is called *substitution away from* an atom  $A$  when  $\sigma$  instantiates each variable  $X$  in  $A$  to a term  $t$  such that every variable in  $t$  is a fresh variable not in  $A$ .

### 2.1. Definition of Data Types

We introduce type into Prolog to separate definite clauses for data structures from others for procedures, e.g.,

```
type.  
  list([ ]).  
  list([X|L]) :- list(L).  
end.
```

**type defines a unary relation by definite clauses. The**

**head of the definite clause takes a term defining a data structure as the argument, either a constant  $b$  called *bottom element* or a term of the form  $c(t_1, t_2, \dots, t_n)$  where  $c$  is called *constructor*. The body shows a type condition about the proper subterms of the argument.**

Here note the set of ground terms prescribed by type predicates. The set of all ground terms  $t$  such that  $!-p(t)$  succeeds is called *type of  $p$*  and denoted by  $\underline{p}$ .

**Example 2.1.** Let the definition of a data type *number* be **type.**

```
number(0).  
number(s(X)) :- number(X).  
end.
```

Then *number* is a set  $\{0, s(0), s(s(0)), \dots\}$ .

Suppose there are defined  $k$  data types  $p_1, p_2, \dots, p_k$  and  $p_1, p_2, \dots, p_k$  are disjoint. We denote the set of all ground terms contained in no  $p_i$  by *others* and consider it like one of types. Then the Herbrand universe  $H$  is divided into  $k + 1$  disjoint sets as follows.

$$H = p_1 \cup p_2 \cup \dots \cup p_k \cup \text{others}.$$

Procedures are defined following the syntax of DEC-10 Prolog [10], e.g.,

```
append([ ], K, K).  
append([X|L], M, [X|N]) :- append(L, M, N).
```

Throughout this paper, we use  $P$  as a finite set of definite clauses defining data types and procedures. We assume variables in each definite clause are renamed at each use so that there occurs no variable names conflict.

### 2.2. A Fundamental Theorem for Type Inference in Prolog

Let  $I$  and  $J$  be Herbrand interpretations.  $I$  is said to *cover success set under a restriction  $J$*  when it contains the intersection of the minimum Herbrand model  $M_0$  and  $J$ .

An Herbrand interpretation  $J$  is said to be *closed with respect to  $P$*  when for any ground instance of definite clause in  $P$  such that the head is in  $J$ , any ground atom in the body is also in  $J$ . This means  $M_0 \cap J$  is computable within it.

**Theorem** If  $J$  is closed with respect to  $P$ ,  $T(I) \subseteq T'(I)$  and  $T'(I) \cap J \subseteq I$ , then  $I$  covers success set under  $J$ .

**Proof.**  $T(I) \cap J \subseteq I$  from  $T(I) \subseteq T'(I)$ . Let  $T_J$  be a monotone transformation of Herbrand interpretations such that  $T_J(I)$  is  $T(I) \cap J$  for any  $I$ .  $T_J$  has a least fixpoint  $\bigcap_{T_J(I) \subseteq I} I$  by the Knaster-Tarski fixpoint theorem ([1] p.843,

Theorem 2.1). Because  $J$  is closed,  $M_0 \cap J$  is a fixpoint of  $T_J$ . Moreover it is the least fixpoint, since  $M_0 \cap J = \bigcup_{i=0}^{\infty} T_J^i(\emptyset)$  ([1] p.843, Theorem 2.2). Therefore  $M_0 \cap J \subseteq I$  for any  $I$  satisfying  $T_J(I) \subseteq I$ .

Our goal of type inference is to describe an Herbrand interpretation  $I$  covering success set under a restriction  $J$  in terms of the types. This is performed by defining an appropriate transformation  $T^t$  satisfying the theorem above.

**3. Type Inference in Prolog**

In this section, we show how to describe a class of Herbrand interpretations in terms of the types first. Then we define an appropriate transformation satisfying the condition in the theorem in 2.2. The basic computation method and the incremental version are presented in 3.4 and 3.5.

**3.1. Interpretation by Type**

(1) Type Set

A set of ground terms represented by a union of types is called *type set*. Type sets are denoted by  $\underline{t}, \underline{t}_1, \underline{t}_2, \dots$  etc.

*Example 3.1.1.* *number*  $\cup$  *list* is a type set.  $\emptyset$  is a type set, too.  $p_1 \cup p_2 \cup \dots \cup p_k$  *Others* is a type set. We denote it by *any*. (*any* is not a type but an abbreviation of a type set.)

(2) Type Substitution

An assignment of type sets to variables

$$\Sigma = \langle X_1 \leftarrow \underline{t}_1, X_2 \leftarrow \underline{t}_2, \dots, X_n \leftarrow \underline{t}_n \rangle$$

is called *type substitution*. A type assigned to a variable  $X$  by  $\Sigma$  is denoted by  $\Sigma(X)$ . We assume  $\Sigma(X) = \text{any}$  for any variable  $X$  not appearing explicitly in the domain of  $\Sigma$ . A type substitution  $\Sigma = \langle X_1 \leftarrow \underline{t}_1, X_2 \leftarrow \underline{t}_2, \dots, X_n \leftarrow \underline{t}_n \rangle$  is considered the same as a set of substitutions

$$\{ \langle X_1 \leftarrow \underline{t}_1, \dots, X_n \leftarrow \underline{t}_n \rangle \mid \underline{t}_i \in \underline{t}_1, \dots, \underline{t}_n \in \underline{t}_n \}.$$

*Example 3.1.2.*  $\langle L \leftarrow \text{list} \rangle$  is a type substitution. This is considered the same as a set of substitution  $\{ \langle L \leftarrow t \rangle \mid t \text{ is any ground term in } \text{list} \}$ . The empty substitution  $\langle \rangle$  is a type substitution assigning *any* to any variable.

The union of two type substitutions  $\Sigma_1$  and  $\Sigma_2$  is a type substitution  $\Sigma$  such that  $\Sigma(X) = \Sigma_1(X) \cup \Sigma_2(X)$  for any  $X$  and denoted by  $\Sigma_1 \cup \Sigma_2$ . The intersection of two type substitutions  $\Sigma_1$  and  $\Sigma_2$  is a type substitution  $\Sigma$  such that  $\Sigma(X) = \Sigma_1(X) \cap \Sigma_2(X)$  for any  $X$  and denoted by  $\Sigma_1 \cap \Sigma_2$ .

(3) Interpretation by Type

Let  $B_0$  be a head of a definite clause " $B_0 :- B_1, B_2, \dots, B_m$ " in  $P$  and  $\Sigma$  be a type substitution. Then  $\Sigma(B_0)$  is considered a set of ground atoms  $\{ \sigma(B_0) \mid \sigma \in \Sigma \}$ . An Herbrand interpretation  $I$  represented by a union of all such forms, i.e.

$$\bigcup_{\langle B_0 :- B_1, B_2, \dots, B_m \rangle \in P} \Sigma(B_0)$$

is called *interpretation by type*.

*Example 3.1.3.* Let  $I$  be an Herbrand interpretation

$$\langle \rangle \text{ (append}(\_, K, K) \cup \langle L \leftarrow \emptyset, M \leftarrow \emptyset, N \leftarrow \emptyset \rangle \text{ (append}(\_, [X|L], M, [X|N])) \text{)}$$

Then  $I$  is an interpretation by type. This is an Herbrand interpretation  $\{ \text{append}(\_, t, t) \mid t \text{ is any ground term} \}$ .

**3.2. Restriction by Type**

An Herbrand interpretation  $J$  of the form  $\bigcup_i \Sigma_i(A_i)$  is called *restriction by type*, where each  $A_i$  is not necessarily a head of definite clauses in  $P$ .

*Example 3.2.*  $J = \langle \rangle \text{ (append}(\_, [A], M) \cup \langle \rangle \text{ (reverse}(\_, M) \text{)}$  is a restriction by type.

**3.3. A Transformation for Type Inference**

(1) Computation of Type Set of Superterm and Subterm

When each variable  $X$  in a term  $t$  is instantiated to a ground term in  $\Sigma(X)$ , we compute a type set containing all ground instances of  $t$  as follows and denote it by  $t/\Sigma$ .

$$t/\Sigma = \begin{cases} \emptyset, & \Sigma(X) = \emptyset \text{ for some } X \text{ in } t; \\ \Sigma(X), & \text{when } t \text{ is a variable } X; \\ p, & \text{when } t \text{ is a bottom element } b \text{ of } p \text{ or} \\ & \text{when } t \text{ is of the form } c(t_1, t_2, \dots, t_n), \\ & c \text{ is a constructor of } p \text{ and} \\ & t_1/\Sigma, t_2/\Sigma, \dots, t_n/\Sigma \text{ satisfy} \\ & \text{the type conditions;} \\ \text{any}, & \text{otherwise.} \end{cases}$$

*Example 3.3.1.* Let  $t$  be  $[X|L]$  and  $\Sigma$  be  $\langle L \leftarrow \text{list} \rangle$ . Then  $t/\Sigma$  is *list*. Let  $t$  be  $[X|L]$  and  $\Sigma$  be  $\langle \rangle$ . Then  $t/\Sigma$  is *any*.

When a term  $t$  containing an occurrence of a variable  $X$  is instantiated to a ground term in  $\underline{t}$ , we compute a type set containing all ground instances of the occurrence of  $X$  as follows and denote it by  $X / \langle t \leftarrow \underline{t} \rangle$ .

$$X / \langle t \leftarrow \underline{t} \rangle = \begin{cases} \underline{t}, & \text{when } t \text{ is a variable } X; \\ X / \langle t_i \leftarrow \underline{t}_i \rangle, & \text{when } t \text{ is of the form} \\ & c(t_1, t_2, \dots, t_n), \underline{t} \text{ is } p, \\ & c \text{ is a constructor of } p, \\ & X \text{ is in } t_i \text{ and} \\ & \underline{t}_i \text{ is a type set assigned} \\ & \text{to the } i\text{-th argument } t_i; \\ \emptyset, & \text{otherwise.} \end{cases}$$

*Example 3.3.2.* Let  $t$  be  $[X|L]$  and  $\underline{t}$  be *list*. Then

$$X / \langle [X|L] \leftarrow \text{list} \rangle = \text{any},$$

$$L / \langle [X|L] \leftarrow \text{list} \rangle = \text{list}.$$

Let  $t$  be  $[X|L]$  and  $\underline{t}$  be *number*. Then

$$X / \langle [X|L] \leftarrow \text{number} \rangle = \emptyset,$$

$$L / \langle [X|L] \leftarrow \text{number} \rangle = \emptyset.$$

(2) Computation of Covering Type Substitution

Let  $B_1, B_2, \dots, B_m$  be a sequence of atoms and  $I$  be an interpretation by type. A type substitution is called *covering type substitution* with respect to  $B_1, B_2, \dots, B_m$  and  $I$  when it contains every substitution  $\sigma$  such that all ground instances  $\sigma(B_1), \sigma(B_2), \dots, \sigma(B_m)$  are in  $I$ . A transformation  $T^t$  is defined using the covering type substitutions.

Let  $I$  be an interpretation by type  $\bigcup_i \Sigma_i(A_i)$  and  $B_1, B_2, \dots, B_m$  be a sequence of atoms. When  $B_1, B_2, \dots, B_m$  are unifiable with  $A_1, A_2, \dots, A_m$  by an mgu  $\tau$ , we define a type substitution  $\Sigma$  on variables in  $B_1, B_2, \dots, B_m$  as follows. Note that we can assume without loss of generality that  $t$  contains no variable in  $B_1, B_2, \dots, B_m$  when a variable  $X$  in  $B_1, B_2, \dots, B_m$  is instantiated to  $t$  by  $\tau$ , because the unifiability shows there is no cycle.

When  $m = 0$  then  $\Sigma = \langle \rangle$ .

When  $m > 0$  then

- (a) Let  $t$  be a term containing variables in  $A_i$  and  $X$  be a variable in  $B_j$ . If  $r$  substitutes  $t$  for  $X$ , then we assigns  $t/\Sigma_i$  to  $X$ .
- (b) Let  $t$  be a term containing an occurrence of  $X$  in  $B_j$  and  $Y$  be a variable in  $A_i$ . If  $r$  substitutes  $t$  for  $Y$ , then we assigns  $X / \langle t \leftarrow \Sigma_i(Y) \rangle$  to the occurrence of  $X$ .

Then  $\Sigma$  assigns the intersection  $\bigcap_i t_i$  to  $X$  when  $t_1, t_2, \dots$  are computed as type sets at different occurrences of  $X$  in  $B_1, B_2, \dots, B_m$ . (When  $\Sigma$  assigns  $\emptyset$  to some variable in  $B_1, B_2, \dots, B_m$ , we assume  $\Sigma$  assigns  $\emptyset$  to any variables.)

By  $[\frac{B_1 B_2 \dots B_m}{\Sigma}]$ , we denote the union of  $\Sigma$  for every possible combination of  $A_i, A_{i_2}, \dots, A_{i_m}$  and its mgu  $r$ .

**Example 3.3.3.** Let  $I$  be a type interpretation

$$\langle \rangle (\text{append}([\ ], K, K)) \cup \\ \langle L \leftarrow \emptyset, M \leftarrow \emptyset, N \leftarrow \emptyset \rangle (\text{append}([X|L], M, [X|N]))$$

and  $B_1$  be a sequence of atoms (though it is only one atom)  $\text{append}(A, B, C)$ .

Then There are two possibility of unification. One is  $\tau_1 = \langle A \leftarrow [], B \leftarrow K, C \leftarrow K \rangle$  and the corresponding type substitution is  $\langle A \leftarrow \text{list} \rangle$ . Another is  $\tau_2 = \langle A \leftarrow [X|L], B \leftarrow M, C \leftarrow [X|N] \rangle$  and  $\langle A \leftarrow \emptyset, B \leftarrow \emptyset, C \leftarrow \emptyset \rangle$  is the corresponding type substitution. Hence by taking their union, we have

$$[\frac{\text{append}(A, B, C)}{\langle A \leftarrow \text{list} \rangle}] = \langle A \leftarrow \text{list} \rangle$$

### (3) A Transformation $T'$

We define  $T'$  as follows.

$$T'(I) = \bigcup_{\langle B_0 \leftarrow B_1, B_2, \dots, B_m \rangle \in P} [\frac{B_1 B_2 \dots B_m}{I}](B_0)$$

It is obvious that  $T(I) \subseteq T'(I)$  and  $T'$  is monotons for interpretations by type.

### 3.4. Computation of Type Inference

An interpretation by type covering success set under a restriction  $J$  is called *type inference under  $J$* . The theorem in 2.2 holds for interpretations by type as well. We already have an appropriate transformation  $T'$  so that we can compute a type inference under a closed restriction  $J$ . The outlook of the basic algorithm for type inference is similar to the bottom-up computation of minimum Herbrand models.

```
i := -1; I0 := ∅;
repeat i := i + 1; Ii+1 := T'(Ii) ∩ J until Ii+1 ⊆ Ii;
return Ii;
```

Figure 1. Computation of Type Inference

In order to compute type inference under a closed restriction  $J$ , we need  $T'(I) \cap J$ , which is obtained by using  $[\frac{B_1 B_2 \dots B_m}{\Sigma}](B_0) \cap [\frac{B_1 B_2 \dots B_m}{\Sigma'}](B_0)$ . ( $[\frac{B_1 B_2 \dots B_m}{\Sigma'}](B_0)$  is common to all repetitions of the type inference process and can be computed once and for all before the repetitions.

### 3.5. Incremental Type Inferences

An atom  $A$  is said to be *closed with respect to  $P$*  when for any ground instance of the definite clause in  $P$  such that the head is a ground instance of  $A$ , any recursive call in the body is also a ground instance of  $A$ . (Hence any

non-recursive definite clause and any definite clause with a head nonunifiable with  $A$  are negligible.) This means the set of ground atoms in  $M_0$  of the form of instance of  $A$  is computable by some instances of definite clauses. Note that  $p(X_1, X_2, \dots, X_n)$  in general form is always closed.

**Example 3.5.1.** An atom  $\text{append}(N, [X], M)$  is closed. This means  $\{\text{append}(t_1, [t_2], t_3) \mid t_1, t_2 \text{ and } t_3 \text{ are ground terms}\} \cap M_0$  is computable by

$$\text{append}([\ ], [Y], [Y]). \\ \text{append}([X|L], [Y], [X|N]) :- \text{append}(L, [Y], N).$$

The closedness of an atom  $A$  can be checked as follows.

- (a) Check whether the head  $B_0$  is unifiable with  $A$  by a substitution for  $A$  away from  $A$  (see section 2). If it is, decompose the mgu to  $\sigma \circ \tau_0$  where  $\sigma$  is the restriction to variables in  $B_0$  and  $\tau_0$  is the restriction to variables in  $A$ . If it is not, neglect the definite clause.
- (b) Check whether each instance of the recursive call in the body  $\sigma(B_i)$  is an instance of  $A$  and if it is, compute the instantiation  $\tau_i$ . If it is not,  $A$  is not closed.

The set of all instances of definite clauses by  $\sigma$  is called *instanciated program for  $A$* .

**Example 3.5.2.** Let the atom  $A$  be  $\text{append}(A, [U], C)$ . Then the first head  $\text{append}([\ ], L, L)$  is unifiable with  $\text{append}(A, [U], C)$  by  $\langle L \leftarrow [Y] \rangle \circ \langle A \leftarrow [], U \leftarrow Y, C \leftarrow [Y] \rangle$ . The second head  $\text{append}([X|L], M, [X|N])$  is unifiable with  $\text{append}(A, [U], C)$  by  $\langle M \leftarrow [Y] \rangle \circ \langle A \leftarrow [X|L], U \leftarrow Y, C \leftarrow [X|N] \rangle$  and the instance  $\text{append}(L, [Y], N)$  in the body is also an instance of  $\text{append}(A, [U], C)$  by  $\langle A \leftarrow L, U \leftarrow Y, C \leftarrow N \rangle$ .

$$\text{append}([\ ], [Y], [Y]). \\ \text{append}([X|L], [Y], [X|N]) :- \text{append}(L, [Y], N).$$

is the instanciated program for  $\text{append}(A, [U], C)$ .

An atom satisfying the following condition is called *closure of  $A$  with respect to  $P$*  and denoted by  $\bar{A}$ .

- (a)  $\bar{A}$  is closed with respect to  $P$ ,
- (b)  $A$  is an instance of  $\bar{A}$  and
- (c)  $\bar{A}$  is an instance of any  $\bar{A}'$  satisfying (a) and (b).

**Example 3.5.3.**  $\text{reverse}(A, B')$  is a closure of  $\text{reverse}(A, [V|B])$ .

The closure is unique up to renaming and  $A$  is closed iff  $A = \bar{A}$  modulo renaming. (See [7] for the proof of uniqueness and the algorithm to compute the closure.)

Now suppose we would like to compute type inference about  $p$  under a restriction by type  $\langle \rangle \overline{p(t_1, t_2, \dots, t_n)}$  and denote it by  $T(\overline{p(t_1, t_2, \dots, t_n)})$ . Let  $A_1, A_2, \dots, A_l$  be non-recursive calls in the bodies of instanciated program for  $\overline{p(t_1, t_2, \dots, t_n)}$ . (If some  $B_i = q(s_1, s_2, \dots, s_m)$  and  $B_j = q(s'_1, s'_2, \dots, s'_m)$ , we distinguish the predicate symbols by  $q_1$  and  $q_2$  and assume both of them have the same definite clause program as  $q$ .) Then we compute  $T(\overline{p(t_1, t_2, \dots, t_n)})$  by initialising  $I_0$  to  $T(\bar{A}_1) \cup T(\bar{A}_2) \cup \dots \cup T(\bar{A}_l)$ , where each  $T(\bar{A}_i)$  is computed recursively.

```
i := -1; I0 := T( $\bar{A}_1$ ) ∪ T( $\bar{A}_2$ ) ∪ ⋯ ∪ T( $\bar{A}_l$ );
repeat i := i + 1; Ii+1 := T'(Ii) until Ii+1 ⊆ Ii;
return Ii - I0;
```

Figure 2. Incremental Type inference

If there is no mutual recursion, this process stops. The base case is the type inference for predicates which call no other predicate, e.g. *append*.

**Example 3.5.4.** Suppose we compute  $T(\text{reverse}(L, M))$ . The computation proceeds by initialising  $I_0$  to  $T(\text{append}(N, [X], M))$  as follows.

$$\begin{aligned} I_0 &= \langle \rangle (\text{append}([], [Y], [Y])) \cup \\ &\quad \langle L \leftarrow \text{list}, N \leftarrow \text{list} \rangle (\text{append}([X|L], [Y], [X|N])) \\ I_1 &= I_0 \cup \langle \rangle (\text{reverse}([], [])) \cup \\ &\quad \langle X \leftarrow \emptyset, L \leftarrow \emptyset, M \leftarrow \emptyset \rangle (\text{reverse}([X|L], M)), \\ I_2 &= I_0 \cup \langle \rangle (\text{reverse}([], [])) \cup \\ &\quad \langle L \leftarrow \text{list}, M \leftarrow \text{list} \rangle (\text{reverse}([X|L], M)) \\ \text{and } I_3 &= I_2. \text{ Hence the type inference of } \text{reverse} \text{ is} \\ &\quad \langle \rangle (\text{reverse}([], [])) \cup \\ &\quad \langle L \leftarrow \text{list}, M \leftarrow \text{list} \rangle (\text{reverse}([X|L], M)) \end{aligned}$$

Recursive computation of  $T(A_i)$  is sometimes unnecessary. For example, when  $A_i$  is an atom  $q(X_1, X_2, \dots, X_m)$  in general form and  $T(q(X_1, X_2, \dots, X_m))$  is already computed before, then recomputing it all the way slows down the whole computation. As another example, when some  $A_i = q(s_1, s_2, \dots, s_m)$  and  $A_j = q(s'_1, s'_2, \dots, s'_m)$ , we distinguish their predicate symbols by  $q_1$  and  $q_2$ . But if  $q(s_1, s_2, \dots, s_m) = q(s'_1, s'_2, \dots, s'_m)$  modulo renaming, it turns out to compute the same result twice and the distinction is useless. In order to avoid the deficiency and accelerate the convergence, we store the results computed before for each closed atom, or more precisely for each instantiated program, and utilize them immediately if possible.

#### 4. Verification of Prolog Programs Using Type Information

Type inference is used effectively in our verification system [5],[6],[7]. In verification we sometimes simplify the theorem to be proved by assuming that some atom is true or false. In such a case, some information about variables left in the simplified theorem may be lost and we need to retain it to prove the right theorem. This problem was first noticed by Boyer and Moore [2] in their theorem prover (BMTP) for pure LISP. The same problem arises in verification of Prolog programs.

**Example 4.** Suppose we prove a theorem

$$\forall U, V, C (\exists B \text{reverse}(C, [V|B]) \supset \exists B' \text{reverse}([U|C], [V|B'])).$$

Because of the definition of *reverse*, it is transformed to

$$\forall U, A, V, C (\exists B \text{reverse}(C, [V|B]) \supset \exists B', D (\text{reverse}(C, D) \wedge \text{append}(D, [U], [V|B'])))$$

Now let us decide  $D$  to be  $[V|B]$ . This decision is sound and we have a new subgoal

$$\forall U, V, C, B (\text{reverse}(C, [V|B]) \supset \exists B' (\text{reverse}(C, [V|B]) \wedge \text{append}([V|B], [U], [V|B'])))$$

Here we can utilize the antecedent. If *reverse*( $C, [V|B]$ ) is false, the theorem is trivially true. Hence we can only need to consider the case *reverse*( $C, [V|B]$ ) is true. By replacing *reverse*( $C, [V|B]$ ) with true, we have a new subgoal

$$\forall U, V, B \exists B' \text{append}([V|B], [U], [V|B'])$$

which is further transformed to

$$\forall U, B \exists B' \text{append}(B, [U], B')$$

because *append*( $[V|B], [U], [V|B']$ ) is *append*( $B, [U], B'$ ). But this transformation has generated a too strong theorem and it is in fact not true. (For example, an instance  $\forall U \exists B' \text{append}(0, [U], B')$  is wrong.) In order to keep the theorem right, we need to add type information as antecedents, i.e. when we derive a subgoal assuming *reverse*( $C, [V|B]$ ) true, we have

information *list*( $B$ ). Our new subgoal should be

$$\forall U, V, B (\text{list}(B) \supset \exists B' \text{append}([V|B], [U], [V|B']))$$

This is provable by induction and we complete the proof.

#### 5. Discussions

Several works are done for type inference in Prolog [3],[8],[9] from different point of views. In our approach, both syntactical and semantical concepts appear. It is semantical whether a type inference / covers a success set, while it is syntactical and closely related to the well-typedness in [3],[9] whether a restriction  $J$  is closed. Moreover they are related strongly through the crucial condition that  $J$  must be closed for / to be computed. Though our approach is still monomorphic, it is new in the following respects.

- Our type inference describes a superset of the success set by associating a type substitutions with each definite clause, which gives some accuracy to the types inferred.
- Our approach solves the problem under a syntactical restriction, which is utilized to infer types more minutely.
- Our type inference is not restricted to that for arguments. Type inference can be done for any variables in any procedure call which is not necessarily in general form.

#### 6. Conclusions

We have shown a type inference method for Prolog programs and its application to verification. This type inference method is an element of our verification system for Prolog programs developed in 1984.

#### Acknowledgements

Our verification system is a subproject of the Fifth Generation Computer System (FGCS) project. The authors would like to thank Dr.K.Fuchi (Director of ICOT) for the chance of this research and Dr.K.Furukawa (Chief of ICOT 2nd Laboratory) and Dr.T.Yokoi (Chief of ICOT 3rd Laboratory) for their advices and encouragements.

#### References

- [1] Apt, K.R. and M.H.van Emden, "Contribution to the Theory of Logic Programming", J.ACM, Vol.29, No.3, pp 841-862, 1982.
- [2] Boyer, R.S. and J.S.Moore, "A Computational Logic", Chap. 6., Academic Press, 1979.
- [3] Bruynooghe, M., "Adding Redundancy to Obtain More Reliable and More Readable Prolog Programs", Proc.1st International Logic Programming Conference, pp.129-133, 1982.
- [4] van Emden, M.H. and R.A.Kowalski, "The Semantics of Predicate Logic as Programming Language", J.ACM, V61.23, No.4, pp.733-742, 1976.
- [5] Kanamori, T., "Verification of Prolog Programs Using An Extension of Execution", ICOT TR-096, 1984.
- [6] Kanamori, T. and H.Fujita, "Formulation of Induction Formulas in Verification of Prolog Programs", ICOT TR-094, 1984.
- [7] Kanamori, T. and K.Horiuchi, "Type Inference in Prolog and Its Applications", ICOT TR-095, 1984.
- [8] Mishra, P., "Towards a Theory of Types in Prolog", Proc. 1984 International Symposium on Logic Programming, pp. 289-298, 1984.
- [9] Mycroft, R. and R.A.O'Keefe, "A Polymorphic Type System for Prolog", Artificial Intelligence 23, pp.295-307, 1984.
- [10] Pereira, L.M., F.C.N.Pereira and D.H.D.Warren, "User's Guide to DECsystem-10 Prolog", Occasional Paper 15, Dept of Artificial Intelligence, Edinburgh, 1979.