知識工学と人工知能 59－4
（１９８８．７．７）

サーカムスクリプションの Reducibilityについて

侯 本慧　富樫 敦　野口 正一
東北大学 電気通信研究所

J. McCarthyによって提案されたサーカムスクリプションは、非単調推論の一種として常識を扱うために非常に有効である。しかし、一階述語のサーカムスクリプションは通常二階の論理式となるため、導出原理などの機械的な証明アルゴリズムでは直接に扱うことができなくなる。本論文では、論理式に対して述語に関するcircumscriptive reducibilityという概念を導入する。これは論理式のサーカムスクリプションと等価な一階の論理式が存在することを意味する。一般的に、任意の論理式はcircumscriptive reducibleであるとは限らない。V. Lifschitzによるseparable formulaがcircumscriptive reducibleである論理式の一例であるが、separable formulasでは知識表現でよく使われているrecursionを表現できない。本論文は、論理式の構文上に制限を加えない形で、circumscriptive reducibleであるための充分条件を提案する。話を簡単にするため、関数記号を含まない論理式を扱っているが、後半で示すように、あるクラスに対しては関数記号を許しても前の結果が成り立つ。

On the Reducibility of Circumscription

Ben-Hui HOU　Atsushi TOGASHI　and　Shoichi NOGUCHI
Research Institute of Electrical Communication –TOHOUKU UNIV.
2－2－1, KATAHIRA SENDAI 980 JAPAN

A form of non-monotonic reasoning, circumscription, has been proposed by J. McCarthy, which possesses the virtue of dealing with commonsense knowledge. However circumscribing first-order predicates makes a given sentence second-order. Then the algorithms of mechanical resolution cannot be directly used in the circumscriptive inference. In this paper, for a theory we introduce a concept of circumscriptive reducibility with respect to a predicate symbol, which says that the result of circumscribing a predicate symbol in a theory has a model-theoretical equivalent counterpart in the same order as that of the theory. Generally, not every theory is circumscriptively reducible. Nevertheless, the theory of the formulas separable with respect to the circumscribed objects, proposed by V. Lifschitz, is circumscriptively reducible. Unfortunately, theories consisting of separable formulas lack the capability to represent recursions on their circumscribed objects. In this paper, we shall propose a condition sufficient to determine the reducibilities of theories no matter whether there are recursions or not. The theories are assumed to be function-free in our discussions. Actually the results are also holding true even if the functions of a large kind are involved.

1. Introduction

Circumscription is a form of non-monotonic reasoning proposed by *J. McCarthy*[3]. for dealing with incomplete knowledge integrally in a conventional logic framework. The original definition will be recalled here.

Let T(P) be a first-order sentence (*i.e.*, a formula without free occurrences of variables) and P a predicate symbol in T. The predicate circumscription of P in T is defined as a second-order formula:

$$T(P) \land \forall P'. [T(P') \land \forall x. (P'(x) \supset P(x)) \supset \forall x. (P(x) \supset P'(x))] \quad (1)$$

where P′ is a predicate variable, T(P′) is the result of substituting P′ for P, and x is a tuple of variables.

The semantics of circumscription is clearer than other non-monotonic logic issues such as default reasoning, while how to draw circumscriptive inference mechanically is relatively difficult because for a theory of first-order, its circumscriptive theory is second-order. However it seems that for some theories, they have corresponding completions, identified with circumscriptions, of the same order as those of original theories themselves. That is, for any formula it can be deduced from circumscriptive theories if and only if it can be deduced from this kind of lower order completions. Here we define such a lower order completion as *circumscriptively reduced* theory and we say a theory is *circumscriptively reducible* when its circumscriptively reduced theory exists. For a special case that the original theory is first-order one, then its circumscriptive inference can be still done without going beyond the framework of first-order logic if it is reducible. However it is not always the case that every theory is reducible. This problem has been mentioned by *J. McCarthy*[4] as "finding the right-substitution for the predicate variables", and he pointed that "Finding the right-substitution for predicate variables, in the cases we have examined, the same task as finding models for a first-order theory.". The circumscriptively reducibility of a theory depends on many aspects. And we have discussed this problem in [1] when the relative theory is restricted within the description of knowledge only with Quosi-IS-A Hierarchy. Here we shall extend our previous result such that it is appropriate to a more general class of theories. Now we shall go into details.

To begin with, we shall reveal the relationship among provability, circumscription and close world assumption. In the coming sections, the main idea of propositions stems from this discussion, and the following expression (2) is very useful to prove our theorems. Let us consider a theory T and a formula F. We introduce a meta-symbol L. L(T, F) stands for that F is provable in theory T. Traditionally, closed world assumption (CWA) [Reiter, R. 1978] suggests that for every formula F, if ¬L(T, F) then ¬F could be admitted as a consequence of T. In other words, what described by T is closed on all predicate symbols. However, circumscribing a predicate symbol P in the theory T means that what described by T is closed only on a certain predicate symbol P. By circumscription, it is possible to make the theory closed on your favorite predicate symbols and for others they still remain traditional logic meaning, while CWA makes the underlying theory closed on all involved predicate symbols. Thus from the view of provability L, circumscription is a more general form of CWA. Using the meta-symbol L, the circumscription of P in T and a theory of T together with CWA could be respectively explained as:

$$T \bigcup \{\forall y.(\forall x. (L(T, P(x)) \land (x \neq y)) \supset \neg P(y))\} \quad (2) \quad \textit{and}$$

$$T \bigcup (\bigcup_{i=1}^{n} \{\forall y_i. (\forall x_i. (L(T, P_i(x_i)) \land (x_i \neq y_i)) \supset \neg P_i(y_i))\}) \quad (3)$$

where all predicate symbols occurring in T are P_1, $P_2, \ldots, P_n$ and $x, x_1, \ldots, x_n, y, y_1, \ldots, y_n$ are tuples of terms.

2. Circumscriptive Reducibility

As for the problem of finding right-substitutions for the predicate variables in the circumscription, there are formulas with the special form, called separable formulas with respect to a predicate symbol, in which circumscribing this predicate symbol could be done in the framework of first-order language. Here we shall recall the concept of separable formulas and point out its weakness is that the theories consisting of separable formulas are lacking capability to cope with recursion through a simple example.

<u>*Definition1*</u>

A formula F is called *solitary* with respect to a predicate symbol P if it is the conjunction of:

(i) formulas without positive occurrences of P;

(ii) formulas of the form $\forall x\, (U(x) \supset P(x))$, where U(x) does not contain P.

A formula F is called *separable* with respect to P if it is a disjunction of solitary formulas.

According to the result by V. *Lifschitz*, the circumscription of a predicate symbol in a formula solitary with respect to this predicate symbol can be represented in the first order language:

$$\text{Circum}(N(P) \wedge U < P; P) = N(U) \wedge (U = P) \qquad (4)$$

where N(P) is the formulas without positive occurrences of P and N(U) is the result of substituting U for P. $U < P$ is a formula of the form $\forall x\ (U(x) \supset P(x))$, $U(x)$ is a formula without occurrences of P. This is the case that relative formula A is solitary with respect to the circumscribed predicate symbol. Similarly, as for the case that the relative formula A is separable with respect to the circumscribed predicate symbol, there are several definitions $U_1, U_2, \ldots, U_n$ of P according to (4) separably. The definition with the minimal extension of P would be the smallest one among $U_1, U_2, \ldots, U_n$, and the smallest U_i is the definition of P obtained by circumscribing P in A. Here we shall not go into details about this and the readers interested in this should be suggested to refer to [6].

As we have seen in the above definition, positive occurrences and negative occurrences of P are not allowed to exist in a formula of the form $\forall x\ (U(x))$ simultaneously. Then formulas recursive on P are obviously not separable. This will be clear in the following example. Now we shall observe this example.

Example1

Let $T \equiv \{ \forall x.\ y.\ (P(a, y) \supset P(x, y)),\ P(a, b),\ P(b, a) \}$. $P(x, y)$ could be considered as the predicate defined by:

$$\forall x.\ y.\ (x = a \wedge y = b) \vee (x = b \wedge y = a) \vee P(a, y) \supset P(x, y).$$

Because there are positive occurrence $P(x, y)$ and negative occurrence $P(a, y)$ of the predicate symbol P in $\forall x.\ y.\ (P(a, y) \supset P(x, y))$, T is not separable with respect to P. However T is circumscriptively reducible with respect to P and the circumscriptively reduced theory is:

$$\forall x.\ y.\ (P(x, y) \equiv (x = a \wedge y = b) \vee (x = b \wedge y = a) \vee (x = b \wedge y = b),$$

if all involved individuals are assumed to be a and b.

Let us consider a formula in the form of $\forall x.\ (Q_P(y) \supset P(x))$, where $Q_P(x)$ is a formula with occurrences of P. Actually it is a recursive formula. By the definition of separable formulas, it is obviously not a separable formula with respect to P. That is, theories consisting of separable formulas lack capability to deal with the formulas with recursiveness. However, as shown in the above example, there are cases in which the first-order circumscription exist even if the relative formulas are not separable with respect to the circumscribing object. Especially when there are no functions involved in the corresponding language, the reducibility of the theory is less dependent on the syntax of involved formulas. In this paper, we shall discuss the reducibility (*i.e.*, there is first-order circumscription *wrt* circumscribed object) of formulas which are not separable *wrt* circumscribed object. In the following discussion no functions are considered. However it will be shown in section6 that there is no much difference when the functions defined by intensional way are involved. The function-freeness in the discussion is just for the sake of convenience.

3. Model-theoretical Meaning of Circumscription

To begin with, several definitions and notations used in this paper are given briefly.

Circum(T; P) is introduced to denote (1), the circumscription [2] of a predicate symbol P on a theory T with the occurrences of P. Here the object involved to be circumscribed is a single predicate symbol P occurring in T. And a restriction is posed on T, say that T consists of only clauses. For each clause Ç in T, $Ç \equiv \ell_1 \vee \ldots \vee \ell_n$, where ℓ_i, $1 \leqq i \leqq n$, are literals, *i.e.*, atomic formulas, or the negation of atomic formulas..

From now on, as we mention a theory we mean it consists of clauses and as we mention a formula we mean it is in clausal form unless different explanations are specially given. For a clause $\ell_1 \vee \ldots \vee \ell_n$, it is logically identified with $\forall x.\ (\ell_1 \vee \ldots \vee \ell_n)$, and x is the tuple of all of variables appearing in this clause.

As for the concept of a circumscriptively reducible theory, we have mentioned it in section1. Here it will be defined formally as follows.

Definition2

Let T be a *first*-order theory consisting of sentences and P a predicate symbol occurring in T. Circum(T; P) is *reducible iff* there is an *first*-order theory, written as $T_{\text{Circum}}(T; P)$, model-theoretically equivalent to Circum(T; P). That is, for any sentence β in the first-order language,

$$\text{Circum}(T; P) \models \beta \quad \textit{iff} \quad T_{\text{Circum}}(T; P) \models \beta.$$

T is said to be *circumscriptively reducible* on P when Circum(T; P) is reducible.

One of the virtues of circumscriptively reducible theories is that it enables mechanical resolution on circumscription of first-order predicates. Obviously, not every theory is circumscriptively reducible. The difficulty that faces us is, what kind of theories are circumscriptively reducible. It will remain to be seen. Now we shall shift onto another definition.

Definition3

The structure M of a sentence A is defined as:
(i) a non-empty Herbrand universe, called the domain of M, denoted by $|M|$;
(ii) $M[K]: |M|^n \rightarrow |M|$, if K is an *n*-ary function symbol;
$M[K]: |M|^n \rightarrow \{True, False\}$ if K is an *n*-ary predicate symbol.

$M[K^+] = \{a \in |M|^n \mid M[K](a) = True\} \subset |M|^n$;

$M[K^-] = \{a \in |M|^n \mid M[K](a) = False\} \subset |M|^n$.

Definition4

Let M and N be two structures of a sentence A. M is a substructure of N in a predicate symbol P, written as $M \leqq_P N$, if
(i) $|M| = |N|$;
(ii) $M[Q] = N[Q]$ for any predicate symbol Q (or function symbols), $Q \neq P$;
(iii) $M[P^+] \subseteq N[P^+]$.

$M[Q^+]$ is called the *extension* of predicate symbol Q in a structure M.

A model M of a sentence A is *minimal in P iff* for any model M′ of A,

$M' \leqq_P M$ *only if* $M' = M$, *i.e.*, $M'[P^+] = M[P^+]$.

Let S be a set of predicate symbols. $M \leqq_S N$ means that $M \leqq_{P_i} N$ for each P_i in S.

P(.) is an atom constructed by an *n*-ary predicate symbol P and a tuple of *n* arbitrary terms, called *P-atom*.

Notice that we divide the evaluation of a predicate symbol into two parts of $M[K^+]$ mapping to {True} and $M[K^-]$ mapping to {False}. And at the definition of $\leqq_P$, only the extensions of P in M and N, $M[P^+]$ and $N[P^+]$, are considered. That is, a model M is said to be smaller than N in P if every predicate symbol except for P is evaluated in the same way and the extension of P in M is a subset of that in N.

Lemma1.1

Let M be a model of T minimal in P. Then for any proper substructure M_0 of M in P, *i.e.*, $M_0 <_P M$, there is at least a formula β containing P in T such that $M_0 \tilde{\ } \vDash \beta$.

【PROOF】 Let M_0 be a proper substructure of M. Then we have

$M_0[K] = M[K]$ for every $K \neq P$ and

$M_0[P] \subset M[P]$, $M_0[P] \neq M[P]$.

M_0 is not a model of T; otherwise we have the contradiction that M is a model of T minimal in P. So that there must be a formula β containing P, such that $M_0 \tilde{\ } \vDash \beta$. **QED**

Lemma1.2

Let M be a model of T. If for any proper substructure M_0 of M, there is at least one β containing P in T such that $M_0 \tilde{\ } \vDash \beta$, then M is a model of T minimal in P.

【PROOF】 Suppose M_0 is a submodel of M for T in P. That is, $M_0 \leqq_P M$. If M_0 is a proper substructure of M in P, then there is β in T, $M_0 \tilde{\ } \vDash \beta$, this contradicts the assumption that M_0 is a model of T. Thus $M_0 = M$. **QED**

Thus according to what suggested in Lemma1.1 and Lemma1.2, to say a model of T is minimal in P is identified with that any substructure of this model is not possibly to be a model of T. Now we shall summarize this as follows:

Proposition1

Let M be a model of T. M is minimal in P *if and only if* for any proper substructure M_0 of M in P, there is at least such a β containing P in T that $M_0 \tilde{\ } \vDash \beta$.

Now according to the above definitions, (1) could be identified with the following formula:

$$T(P) \wedge \neg \exists P'. (T(P') \wedge \forall x. (P'(x) \supset P(x) \wedge P' \neq P) \quad (1')$$

Observing the the definition of minimal model and (1′), we can figure out that the structures of sentence A satisfying (1′) are models of A minimal in P. Actually models of A minimal in P are the only ones satisfying (1′). This will be mentioned in section5.

Readers interested in details on the definitions of structure, model and minimal model, *etc.*, are suggested to refer to the book written by Joseph R. Shoenfield [5].

4. Operator Γ and Π_P

Let T be a theory in clausal form, that is, a set of clauses of the form $\ell_1 \vee \ldots \vee \ell_n$. **Res(T)** is the set of all those resolvents derivable from any pair of clauses in T. It will be detailed below.

Notation [Ç] will be used to describe **Res(T)**. [Ç] is the set of all literals in Ç and identified with $\{\ell_1, \ldots, \ell_n\}$ if $Ç = \ell_1 \vee \ldots \vee \ell_n$. Ç is the disjunction of all elements in [Ç] and identified with $\ell_1 \vee \ldots \vee \ell_n$ if $[Ç] = \{\ell_1, \ldots, \ell_n\}$.

For any clauses $Ç_i$, and $Ç_j$, $Ç_i = \ell_1 \vee \ldots \vee \ell_n$, $Ç_j = \gamma_1 \vee \ldots \vee \gamma_m$ in T,

$[Ç_iÇ_j] = ([Ç_i] - \ell_i)\delta \cup ([Ç_j] - \gamma_j)\delta$,
$Ç_iÇ_j \in$ **Res**(T),
if there is a substitution δ such that $\ell_i\delta = \neg\gamma_j\delta$, and $\ell_i \in [Ç_i]$, $\gamma_j \in [Ç_j]$.

Now we shall introduce an operator Γ based on the definition of **Res**. For a theory T, it will be extended by applying **Res** and for this extended one, it may be extended furthermore in the same way. Take all of those theories as Γ(T), then Γ(T) is a set of all formulas provable in T when Γ(T) cannot be extended by **Res** any more. This will be stated after the definition of Γ has been formally given.

Definition5

Let T be a theory in clausal form. Γ is an operator on T, defined as:
(i) $\Gamma^{(0)}(T) = \mathbf{Res}(T) \cup T$;
(ii) $\Gamma^{(i+1)}(T) = \Gamma^{(i)} \cup \mathbf{Res}(\Gamma^{(i)})$
Then we define:

$$\Gamma(T) = \bigcup_{i=0}^{\infty} \Gamma^{(i)}(T).$$

Γ(T) is denoted by Π(T) when Γ(T) is finite. Now we shall go to observe Γ. Firstly Γ(T) cannot be extended by **Res** any more.

Proposition2 $\Gamma((\Gamma(T)) = \Gamma(T)$.

【PROOF】 It is obvious that

$$\mathbf{Res}\,(\bigcup_{i=0}^{\infty} \Gamma^{(i)}(T)) \supseteq \bigcup_{i=0}^{\infty} \Gamma^{(i)}(T).$$

And for any $\Gamma^{(i)}(T)$, we have,

$$\mathbf{Res}\,(\Gamma^{(i)}(T)) \subseteq \bigcup_{i=0}^{\infty} \Gamma^{(i)}(T).$$

Then

$$\bigcup_{i=0}^{\infty} Res(\Gamma^{(i)}(T))) \subseteq \bigcup_{i=0}^{\infty} \Gamma^{(i)}(T).$$

Because of
$\mathbf{Res}(\Gamma^{(0)}(T)) \cup \mathbf{Res}(\Gamma^{(1)}(T)) \cup \ldots =$
$\mathbf{Res}(\Gamma^{(0)}(T) \cup \Gamma^{(1)}(T) \cup \ldots)$,

$$\mathbf{Res}\,(\bigcup_{i=0}^{\infty} \Gamma^{(i)}(T)) \subseteq \bigcup_{i=0}^{\infty} \Gamma^{(i)}(T).$$

Thus $\mathbf{Res}(\Gamma(T)) = \Gamma(T)$. **QED**

Secondly, Γ(T) is finite when T is a theory without occurrences of functions.

Proposition3

There is such an integer n that

$$\Gamma(T) = \bigcup_{i=0}^{n} \Gamma^{(i)}(T)$$

and $\Gamma(\Gamma(T)) = \Gamma(T)$ if T is function-free.

【PROOF】 Suppose there is no such an n,

$$\Gamma(T) = \bigcup_{i=0}^{n} \Gamma^{(i)}(T)$$

that $\mathbf{Res}(\Gamma(T)) = \Gamma(T)$. That is for any k, $\Gamma^{(k)}(T) \subset \Gamma^{(k+1)}(T)$.

However, T is function-free and T is a finite set. Then the set of all clauses available from T (*i.e.*, constructed by predicate symbols and constants occurring in T) is finite, denoted by EM[T]. Obviously $\Gamma^{(k)}(T) \subseteq EM[T]$.

This contradicts with the supposition. Thus there is an n,

$$\Gamma(T) = \bigcup_{i=0}^{n} \Gamma^{(i)}(T)$$

such that $\mathbf{Res}(\Gamma(T)) = \Gamma(T)$. **QED**

Thirdly, Γ(T) is the set of all formulas provable in T when Γ(T) is finite.

Proposition4

Let β be a formula and T a theory. β is provable from T *iff* β contains an instance of some formula α in Π(T).

Corollary4.1

Let β be a formula of the form P(x) and T a theory of clausal form. β is provable from T *iff* β is an instance of β′, for some $\beta' \in \Pi_P(T)$.

And at last, we introduce a notation $\Pi_P(T)$ to represent the set of all P-atoms in Π(T).

Definition6

$\Pi_P(T) = \{P(\,.\,) \mid \text{for any } P(\,.\,) \text{ in } \Pi(T)\}$.

Based on all of those definitions and propositions we shall go to our main issue of this paper.

5. Reducibility on Γ

We shall see that the reduced theory $T_{Circum}(T; P)$ corresponding to a circumscriptively reducible theory T with respect to P contains following equality axioms and unique name hypothesis.

Equality axioms (EA) are:
(EA1) $\forall x.\, x = x$
(EA2) $\forall x\ y.\, x = y \supset y = x$
(EA3) $\forall x\ y\ z.\, x = y \wedge y = z \supset x = z$
(EA4) For each n-ary predicate symbol P,
$\forall x_1, \ldots, x_n;\ y_1, \ldots, y_n.$
$x_1 = y_1 \wedge \ldots \wedge x_n = y_n \wedge P(x_1, \ldots, x_n)$
$\supset P(y_1, \ldots, y_n)$.

Unique name hypothesis (UNH) indicates that: the assumption that each object has a unique name, *i.e.*, that distinct names denote distinct objects, written as $x \neq y$, says that x and y are two distinct names. Here we say '=' *equality* and '≠' *distinctness*.

Equality axioms and unique name hypothesis declare that in $T_{Circum}(T; P)$, two constants are distinct unless their equality has been proved from (EA1)~(EA4). Thus $T_{Circum}(T; P)$ suggests $\neg P(c)$ if for any $P(c_i)$ provable in T, the distinctness of c and c_i has been proved.

Now we introduce two concepts of *circumscriptive inference*, denoted by '$\vdash_P$', and *minimal entailment*, denoted by '$\models_P$'. Let T be a theory and β a formula. β is said to circumscriptively inferred from T with respect to P, written as $T \vdash_P \beta$, when it is inferred from the result of T by circumscribing P. And β is said to be minimally entailed from T with respect to P, written as $T \models_P \beta$, when it is true in every model of T minimal in P.

It has been known that the circumscriptive inference is sound in term of minimal entailment. That is, all of formulas circumscriptively inferred from T can be minimally entailed from T.

<u>Theorem1</u>[2],[6]

If $T \vdash_P \beta$ then $T \models_P \beta$.

However above theorem cannot always be conversed because for a theory T, not every model of T has a model minimal in P.

Let the relative theory T be of definite clausal. Then every model of T has an identified minimal model. In this case, Π(T) is the counterpart of this minimal model. The theory of T and all of enumerations of $\Pi_P(T)$ together with equality axioms and unique names hypothesis is actually $T_{Circum}(T; P)$, the reduced circumscription of P in T. This is formally shown and proved below under the condition that T is function-free. It is only for the sake of convenience. It will be seen that there is no much difference when functions of certain kind are involved. We shall firstly consider (5) as the theory of T and all of enumerations of $\Pi_P(T)$ preparatory to proposing following theorems.

$$T_{Circum}(T; P) \equiv T \cup \{\forall. x_1, \ldots, x_n\, (P(x_1, \ldots, x_n) \equiv \bigvee_{1 \leq i \leq r} \exists y_{i1}, y_{i2}, \ldots, y_{im_i} (x_1 = t_{i1} \wedge x_2 = t_{i2} \wedge \ldots \wedge x_n = t_{in})\} \quad (5)$$

For any i, $1 \leqq i \leqq r$, $y_{i1}, y_{i2}, \ldots, y_{im_i}$ are variables in $P(t_{i1}, t_{i2}, \ldots, t_{in})$, where $P(t_{i1}, t_{i2}, \ldots, t_{in}) \in \Pi_P(T)$, and r is the number of P-atoms in $\Pi_P(T)$.

Let $x = \{x_1, \ldots, x_n\}$ $y_i = \{y_{i1}, y_{i2}, \ldots, y_{im_i}\}$, and $t_i = \{t_{i1}, \ldots, t_{in}\}$.

$\exists y_i (x = t_i)$ stands for

$\exists y_{i1}, y_{i2}, \ldots, y_{im_i} (x_1 = t_{i1} \wedge x_2 = t_{i2} \wedge \ldots \wedge x_n = t_{in})$.

Then $\exists y_i (x \neq t_i)$ means that there is at least one j, $1 \leqq j \leqq n$, such that $x_j \neq t_{ij}$.

$\vee \exists y_i (x = t_i)$ stands for

$\bigvee_{1 \leq i \leq r} \exists y_{i1}, y_{i2}, \ldots, y_{im_i}$
$(x_1 = t_{i1} \wedge x_2 = t_{i2} \wedge \ldots \wedge x_n = t_{in})$.

$\wedge \exists y_i (x \neq t_i)$ means that, for every i, $1 \leqq i \leqq r$, $x \neq t_i$.

By all of those notations and assumption, (5) can be written as:

$$T_{Circum}(T; P) \equiv T \cup \{\forall x. [P(x) \equiv (\exists y_1 (x = t_1) \vee \exists y_2 (x = t_2) \vee \ldots \vee \exists y_r (x = t_r))]\}.$$

Furthermore it can be simplified as:

$$T \cup \{\forall x. (P(x) \equiv \vee \exists y_i (x = t_i))\} \quad (5')$$

As we shall see later, (5′) together with EA and UNH is $T_{Circum}(T; P)$, the reduced circumscription of P in T.

When the set of P-atoms true in T does not vary with the different models of T, together with what suggested by Lemma3 we have the following theorem.

<u>Theorem2</u>

T is circumscriptively reducible on any predicate symbol in T if T is a definite clausal theory without functions.

【PROOF】 We take (5) as $T_{Circum}(T; P)$.

We shall prove the theorem by the proofs of:

(i) If $Circum(T; P) \models \beta$, *i.e.*, $T \models_P \beta$, then $T_{Circum}(T; P) \models \beta$; and

(ii) If $T_{Circum}(T; P) \models \beta$, then $T \models_P \beta$, *i.e.*, $Circum(T; P) \models \beta$.

<u>Proof of (i)</u>:

Suppose M_1 is a model for $T_{Circum}(T; P)$. We will show M_1 is a model of T minimal in P.

(a) M_1 is obviously a model of T;

(b) Prove M_1 is minimal in P.

Now assume M_2 is any model of T with $M_2 \leqq_P M_1$. Then it is sufficient to show $M_1[P^+] = M_2[P^+]$ in order to prove that M_1 is minimal in P.

By $M_2 \leqq_P M_1$, we have

$M_2[K] = M_1[K]$ if $K \neq P$ and

$M_2[P^+] \subseteq M_1[P^+]$.

For any $P(t_{i1}, \ldots, t_{in})$ true in M_1, *i.e.*, $(t_{i1}, \ldots, t_{in}) \in M_1[P^+]$, there is at least one $P(t_{i1}', \ldots, t_{in}')$ in $\Pi_p(T)$ such that $P(t_{i1}, \ldots, t_{in})$ is an instance of

$P(t_{i1}', ..., t_{in}')$, because $\forall x.(P(x) \supset \bigvee \exists y_i\ (x=t_i))$ in $T_{Circum}(T; P)$ must be satisfied by M_1. $P(t_{i1}, ..., t_{in})$ is provable from T by Corollary4.1, *i.e.*, $T \vdash P(t_{i1}, ..., t_{in})$. By the soundness of $\vdash$, we have $T \models P(t_{i1}, ..., t_{in})$. Thus $P(t_{i1}, ..., t_{in})$ is true in M_2 because M_2 is a model of T. Then we can say $(t_{i1}, ..., t_{in}) \in M_2[P^+]$. Hence $M_1[P^+] \subseteq M_2[P^+]$. Together with $M_2[P^+] \subseteq M_1[P^+]$, we get $M_2[P^+] = M_1[P^+]$. Therefore, the model M_1 of $T_{Circum}(T; P)$ is also a model of T minimal in P. That is, if $Circum(T; P) \models \beta$, then $T_{Circum}(T; P) \models \beta$.

Proof of (ii):

We shall prove that if $T_{Circum}(T; P) \models \beta$ then $Circum(T; P) \models \beta$, *i.e.*, $T \models_P \beta$.

Suppose M_0 be a model of T minimal in P, i.e., for any model M of T with $M \leqq_P M_0$, we have $M = M_0$, *i.e.*,

$M_0[K] = M[K]$ for every K.

It is sufficient to show $\forall x.\ (\bigvee \exists y_i\ (x=t_i) \supset P(x))$ and $\forall x.(P(x) \supset \bigvee \exists y_i\ (x=t_i))$ in $T_{Circum}(T; P)$ are true in M_0 in order to show M_0 is a model of $T_{Circum}(T; P)$ because M_0 is a model of T.

(a) To show $\forall x.\ (\bigvee \exists y_i\ (x=t_i) \supset P(x))$ in $T_{Circum}(T; P)$ is true in M_0. Let $t_{i1}', ..., t_{in}'$ be any terms satisfying $\exists y_{i1}, ..., y_{im_i}.\ (t_{i1}'=t_{i1} \wedge \ldots \wedge t_{in}'=t_{in})$. Then $P(t_{i1}', ..., t_{in}')$ is an instance of $P(t_{i1}, ..., t_{in})$ in $\Pi_P(T)$ according to (5). By Corollary4.1, $P(t_{i1}', ..., t_{in}')$ is provable from T. Then $P(t_{i1}', ..., t_{in}')$ is true in M_0 because M_0 is a model of T. That is, $\forall x.\ (\bigvee \exists y_i\ (x=t_i) \supset P(x))$ is true in M_0;

(b) To show $\forall x.(P(x) \supset \bigvee \exists y_i\ (x=t_i))$ in $T_{Circum}(T; P)$ is true in M_0. Let $P(t_{i1}', ..., t_{in}')$ be true in M_0. Suppose $P(t_{i1}', ..., t_{in}')$ is not an instance of any $P(t_{i1}, ..., t_{in})$ in $\Pi_P(T)$. Then we can construct a proper substructure M of M_0 in the following way:

$M[K] = M_0[K]$ if $K \neq P$ and

$M[P^+] = M_0[P^+] - (t_{i1}', ..., t_{in}')$.

M is obviously a model of T with $M <_P M_0$. This contradicts the minimality of M_0. Hence M_0 satisfies $\forall x.(P(x) \supset \bigvee \exists y_i\ (x=t_i))$.

Then the minimal model M_0 of T in P is a model of $T_{Circum}(T; P)$ also. That is, if $T_{Circum}(T; P) \models \beta$ then $Circum(T; P) \models \beta$, *i.e.*, $T \models_P \beta$.

Thus $T_{Circum}(T; P)$ is actually a circumscriptively reduced theory of T *wrt* P and T is circumscriptively reducible *wrt* P. **QED**

As mentioned above, for a theory of definite clausal form, it is circumscriptively reducible by virtue of the uniqueness of its minimal model. What would be the case that not only definite clauses but also indefinite ones are involved? Evidently, it dose not take the advantage of the uniqueness of minimal model any more. Then $\Pi(T)$ is not still a counterpart of any one of those minimal models and $\Pi_P(T)$ does not correspond to the extension of P in any one of those minimal models for T. Let us examine an indefinite clause $Q(x) \vee P(x)$. It has at most two minimal models of satisfying $Q(x)$ and $P(x)$ respectively. Now take every minimal model of $Q(x) \vee P(x)$ in such a way that $\neg Q(x)$ could be satisfied. Then $Q(x) \vee P(x)$ has only one minimal model. However this is not always the case. The problem encountered here is that there may be some minimal models of $Q(x) \vee P(x)$ in which $\neg Q(x)$ could not be satisfied. Anyway, $Q(x) \vee P(x)$ has a unique minimal model if for any minimal model of $Q(x) \vee P(x)$, $\neg Q(x)$ is always true. Hinted by this, we could figure out that when $\Pi_P(T)$ could not be expanded by adding arbitrary formulas, then $\Pi_P(T)$ becomes a counterpart of the extension of P in every minimal model for T even if there is no minimal model of T corresponding to $\Pi(T)$.

Theorem3

Let T be a theory of clausal form without functions.

Circum(T; P) is reducible if there is no $\tilde{N}_P$ such that

$\beta \in \Pi_P(T \cup \tilde{N}_P) - \Pi_P(T)$ and $\beta \notin \Pi_P(\tilde{N}_P)$,

for some P-atom β.

The circumscriptively reduced theory of T on P, $T_{Circum}(T; P)$, is

$$T \cup \{\forall.\ x_1, \ldots, x_n\ (P(x_1, \ldots, x_n) \equiv \bigvee_{1 \leq i \leq r} \exists y_{i1}, y_{i2}, \ldots, y_{im_i}\ (x_1 = t_{i1} \wedge x_2 = t_{i2} \wedge \ldots \wedge x_n = t_{in})\} \quad (5)$$

For any i, $1 \leqq i \leqq r$, $y_{i1}, y_{i2}, ..., y_{im_i}$ are variables in $P(t_{i1}, t_{i2}, ..., t_{in})$, where $P(t_{i1}, t_{i2}, ..., t_{in}) \in \Pi_P(T)$, and r is the number of P-atoms in $\Pi_P(T)$.

【PROOF】 At first, we shall give the outline of the proof. Next, we will go into details.

To begin with, it can be shown that every model of $T_{Circum}(T; P)$ is a model of T minimal in P, *i.e*, $T_{Circum}(T; P) \models \beta$ if $Circum(T; P) \models \beta$, *i.e*, $T \models_P \beta$.

Then show that every model of T minimal in P is a model of $T_{Circum}(T; P)$, *i.e.*, $Circum(T; P) \models \beta$ if $T_{Circum}(T; P) \models \beta$, when the condition presented in this theorem is satisfied by T.

(a) Show $T_{Circum}(T; P) \models \beta$ if $Circum(T; P) \models \beta$.

Suppose M_0 is a model of $T_{Circum}(T; P)$. Obviously it is a model of T. Now we shall prove that M_0 is minimal on P. That is, for any model M of T with $M \leqq_P M_0$, we have $M_0[P^+] = M[P^+]$.

By $M \leqq_P M_0$, we have:

$M[K] = M_0[K]$ if $K \neq P$ and

$M[P^+] \subseteq M_0[P^+]$.

For any $P(t_{i1}, ..., t_{in})$ true in M_0, *i.e.*, $(t_{i1}, ..., t_{in}) \in M_0[P^+]$, there is at least one $P(t_{i1}', ..., t_{in}')$ in $\Pi_P(T)$ such that $P(t_{i1}, ..., t_{in})$ is an instance of

$P(t_{i1}', ..., t_{in}')$, because $\forall x.(P(x) \supset \vee \exists y_i\ (x=t_i))$ in $T_{Circum}(T; P)$ must be satisfied by M_0. $P(t_{i1}, ..., t_{in})$ is provable from T by Corollary4.1, *i.e.*, $T \vdash P(t_{i1}, ..., t_{in})$. By the soundness of $\vdash$, we have $T \models P(t_{i1}, ..., t_{in})$. Thus $P(t_{i1}, ..., t_{in})$ is true in M because M is a model of T. Then we can say $(t_{i1}, ..., t_{in}) \in M[P^+]$. Hence $M_0[P^+] \subseteq M[P^+]$. Together with $M[P^+] \subseteq M_0[P^+]$, we get $M[P^+] = M_0[P^+]$. Therefore, the model M_0 of $T_{Circum}(T; P)$ is a model of T minimal in P. That is, if $Circum(T; P) \models \beta$, then $T_{Circum}(T; P) \models \beta$.

(b) Show $Circum(T; P) \models \beta$ if $T_{Circum}(T; P) \models \beta$.

Now we shall prove that any model M of T minimal on P is a model of $T_{Circum}(T; P)$.

Suppose some model M of T minimal in P is not a model of $T_{Circum}(T; P)$. Then it must be that $\forall x.\ (P(x) \supset \vee \exists y_i.\ (x=t_i))$ is not satisfied by M because M is a model of T and $\forall x.\ (\vee \exists y_i.\ (x=t_i) \supset P(x))$ is true in M, which can be shown similarly as in the proof of theorem2. That is, there is at least one $P(t_{i1}', ..., t_{in}')$ satisfied by M but $P(t_{i1}, ..., t_{in})$ in $\Pi_P(T)$ such that $\exists y_{i1}, ..., y_{im_i}.\ (t_{i1}'=t_{i1} \wedge ... \wedge t_{in}'=t_{in})$, cannot be satisfied by M. Then we construct a substructure M_0 of M in the way:

$M_0[K] = M[K]$ if $K \neq P$ and

$M_0[P] = M[P] - (t_{i1}', ..., t_{in}')$.

We shall then distinguish two cases.

Case1. When no formula of the form $P(t_{i1}', ..., t_{in}') \vee Q'(x)$, which is an instance of some formula in $\Pi(T)$, is falsified by M_0, where $Q'(x)$ may be any disjunction of literals. Obviously, M_0 is a model of T and $M_0 < M$. This contradicts the minimality of M;

Case2. When some formula of the form $P(t_{i1}', ..., t_{in}') \vee Q'(x)$, which is an instance of some formula $P(t_{i1}, ..., t_{in}) \vee Q(x)$ in $\Pi(T)$, is falsified by M_0. That is, $Q'(x)$ is false in M_0. Then take $\{\neg Q(x)\}$ as $\tilde{N}_P$. We have $P(t_{i1}, ..., t_{in}) \in \Pi_P(T \cup \tilde{N}_P) - \Pi_P(T)$ and $P(t_{i1}, ..., t_{in}) \notin \Pi_P(\tilde{N}_P)$ because $P(t_{i1}, ..., t_{in})$ is not satisfied by the model M of T (*i.e.*, $P(t_{i1}', ..., t_{in}') \notin \Pi_P(T)$). This contradicts the given condition of our theorem.

Then the minimal model M of T in P is a model of $T_{Circum}(T; P)$ when the condition presented in the theorem is satisfied by T. Hence $Circum(T; P) \models \beta$ if $T_{Circum}(T; P) \models \beta$. Therefore we have proven that $T_{Circum}(T; P)$ is a circumscriptively reduced theory of T and T is circumscriptively reducible *wrt* P. **QED**

This theorem suggests that there is a kind of circumscriptively reducible theories with recursive formulas (of course, not separable). This can also be seen in Example1 in the section2. Recall we have mentioned that the condition of function-freeness placed upon the relative theories is not very strict because the involvement of a large kind of functions usually encountered in knowledge bases dose not cause a lot of trouble in maintaining previous theorems. Now let us go into this issue: recursiveness and functions.

6. Recursiveness and Functions

As discussed in the section2 the class of separable formulas lacks the capability of dealing recursiveness, which is very important in representing knowledge with logic language. And also the previous discussion has been done under the condition of that the relative formulas are function-free. Now we will show that in the field of knowledge base the function-freeness is not a very strong condition. The results are still holding when functions of certain kind are involved.

Suppose f be an n-ary function symbol. f is defined as:

$f: D_1 \times D_2 \times \ldots \times D_n \rightarrow D_0$ and

for any i, $0 \leqq i \leqq n$, D_i is a set of individuals.

Firstly, we introduce an $n+1$-ary predicate symbol F, defined as:

$F(x_1, x_2, \ldots, x_n, y)$ is true *iff* $f(x_1, x_2, \ldots, x_n) = y$.

That is, for a formula $Q(f(x_1, \ldots, x_n))$ with the occurrences of f, it can be substituted by $Q(y) \wedge F(x_1, \ldots, x_n, y)$ and F is defined by a set of atomic formulas of $F(x_1, \ldots, x_n, y)$ *iff* there is $f(x_1, \ldots, x_n) = y$. When all of $D_0, D_1, \ldots, D_n$ is finite f is said to be *representable* by the predicate symbol F.

Let us consider the case that every D_i is finite. Then a function representable by a predicate symbol F corresponds to a finite number of F-atoms. Take a theory with functions. If all of those functions are representable by predicate symbols, a function-free theory, the result of the theory substituting all of the occurrences of functions with their corresponding substitutors of atoms, is identical with the original one.Obviously there is no trouble to extend our previous results by involving representable functions. At the same time, we also know, in databases or knowledge bases there are no much functions which are not representable ones.

Example2

Let $T \equiv isblock(A) \vee isblock(B)$. Suppose $N_P = \neg isblock(A)$. Then we have:

$\Pi_P(T \cup \Pi_P) - \Pi_P(T)$

$= \Pi_P(\{isblock(A) \vee isblock(B)\} \cup \{\neg isblock(A)\}) - \Pi_P(\neg isblock(A))$

$= \{isblock(B)\}$ and

$isblock(B) \notin \Pi_P(N_P)$ $(\Pi_P(N_P) = \emptyset)$.

Thus whether $isblock(A) \vee isblock(B)$ is reducible on isblock cannot be determined by the condition presented in our paper.

According to above two examples, we see that both separability of a theory and the condition in the coming proposition are very strong. There are theories whose reducibility can be determined by the proposition, while they are not separable, shown in example1. There are also separable theories whose reducibility cannot be determined by the condition presented in the proposition, shown in example2.

7. Conclusion

Now we conclude with the comparison among several concepts:

— Every theory in definite clausal form, *i.e.*, each formula in the theory is $\ell_1 \vee \ldots \vee \ell_n$, $n \leqq 1$, is reducible because the condition appearing in Theorem3 is always satisfied by a definite clausal theory;

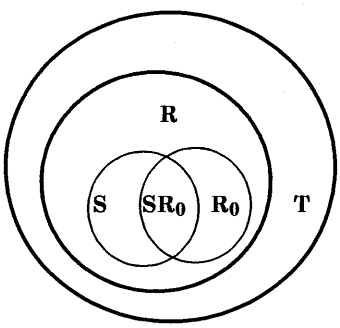
— T(P) is separable *wrt* P *iff*[5] T(P) consists of:

(1) formulas without positive occurrences of P;

(2) formulas of the form $\forall x\ (F(x) \supset P(x))$, where $F(x)$ does not contain P.

And we also know that a theory is circumscriptively reducible on a predicate symbol if it is separable on this predicate symbol.

As shown in the example1 and example2, the class of reducible theories determined by the proposed condition here is joint with the class of theories separable on certain predicate symbol. This can be represented by the following figure:



T is the set of all considered theories;
R is the set of all reducible theories;
S is the set of all separable theories;
$\mathbf{R_0}$ is the set of all reducible theories determined by the proposition.

— What suggested here is also just a condition sufficient for the reducibility of a theory. Thus the converse of Theorem3 is not always holding.

References

[1] B. H. Hou, A. Togashi and S. Noguchi *The Representation of the Knowledge with Hierarchy* 情報処理学会第35回(昭和62年後期)全国大会P1905－1906

[2] B. H. Hou, A. Togashi and S. Noguchi *Circumscription Schema*のオーダー の縮小について 情報処理学会第36回(昭和63年前期)全国大会 P1339－1340

[3] John McCarthy *Circumscription — A Form of Non-Monotonic Reasoning* AI 13(1980) 27–39

[4] John McCarthy *Applications of Circumscription to Formalizing Common Sense Knowledge* AI 28(1986) 89–116

[5] Joseph R.SHOENFIELD Mathematical Logic

[6] Lifschitz, V., *Computing Circumscription* Proceedings Ninth International Joint Conference on AI, Los Angeles, CA (1985) 121–127

[7] Martin Davis *The Mathematics of Non-Monotonic Reasoning* AI 13 (1980) 73–80

[8] Reiter, R. [1978] *On Closed World Data Bases* In Logic and Databases, H. Gallaire and J. Minker, Eds. Plenum, New York, pp 56–76

[9] Reiter, R. [1980] *Equality and Domain Closure in First-Order Databases* JACM Vol. 27, No. 2 pp 235–249