TR-315

Ascription: Application of the Circumscription Technique to Various Conjectural Reasoning

> by J. Arima

November, 1987

©1987, ICOT



Mita Kokusai Bldg. 21F 4-28 Mita 1-Chome Minato-ku Tokyo 108 Japan

(03) 456-3191 ~ 5 Telex ICOT J32964

Ascription: Application of the Circumscription Technique to Various Conjectural Reasoning

Jun Arima

ICOT Research Center
Institute for New Generation Computer Technology
Mita Kokusai Bldg. 21F
4-28 Mita 1-chome, Minato-ku, Tokyo, 108, Japan

Phone: +81 3 456 4365, C.Mail Address: arima@icot.junet or arima%icot.jp@csnet-relay

Abstract

Under the circumstances in which it is always possible to introduce new axioms, any reasoning which jumps to conclusions will be forced to have a non-monotonic property to maintain consistency. In the real world, induction and analogy as well as common sense reasoning are no exceptions.

This paper attempts to formalize such conjectural reasoning processes uniformly as non-monotonic reasoning. For this purpose, the *circumscription* [6,7] technique seems to be most hopeful, especially *formula circumscription* [7]. Unfortunately, formula circumscription is too general. This is shown by presenting an instance which seems too general.

This paper proposes a logical framework, called ascription, for the above purpose. Ascription is realized using the circumscription technique, based on predicate substitution, and its preservation of consistency is guaranteed. Its formalization is not one which makes it possible to perform such conjectural reasoning mechanically, but one which is expected to clarify future research.

1. Introduction

Computer systems with capabilities of deductive inference will release man from the troublesome task of procedural programming and be able to solve problems which are given only declaratively. However, since deductive inference is the deduction of properties of individuals from given general knowledge, it cannot provide effective consequences about facts that are unexpected and not included in general knowledge. This means that deductive inference cannot make a significant contribution to solving our software crisis.

We should remember that unexpected facts always exist and that deductive inference is helpless with respect to them. One promising approach is reasoning by relativizing and generalizing acquired knowledge so that it can be applied to unexpected circumstances.

Research related to this kind of reasoning has been done by McCarthy. Circumscription [6,7] is a form of conjectural reasoning done by humans and based on the closed-world assumption. This work is important and interesting, but in the case of only minimizing the extension of some predicate symbols, it seems that it can explain only a small part of human flexible reasoning and that there still remain some very important aspects which we should not ignore. These are analogy, induction and other reasoning processes which strongly relativize and generalize knowledge. Such reasoning is closely related to human learning capabilities. We have studied reasoning from this point of view and propose a logical framework, called ascription, which is a form of such conjectural reasoning. Intuitively, ascription represents the flexible notion that the interpretation of a certain property K lies between two extremes; one, similar to predicate circumscription, that the only demonstrated positive instances of K are all instances satisfying K; and the other, that all but the demonstrated negative instances satisfy K. More precisely, we will show this as follows.

Ascription is based on the following notion: if any model of a set of formula, Γ , can be transformed into an other model of Γ in which all the entities satisfying a property, K, also satisfy a property, Ψ , by reducing the extension of K to the intersection of the extensions of K and Ψ , and any model of Γ can also be transformed into another model of Γ in which all the entities not satisfying K also do not satisfy Ψ by reducing the extension of K to the union of the extensions of K and Ψ , it may well be considered that K is equivalent to Ψ . That is, when all the demonstrated positive instances of K are positive instances of K, and similarly all the demonstrated negative instances of K are negative instances of K, we can assume the equivalence of K and Ψ .

2. To Formalize Various Types of Conjectural Reasoning

The original concept of circumscription [6] was like the concept of Occam's razor, in that the only entities which are forced to have a certain property by given knowledge are all those that have the property. This concept is very clear and suited to various knowledge racing fields which are based on the closed-world assumption and which have succeeded in practice. However, when we consider a computer system as an intelligent tool, something like a thinking or learning machine, this concept is weak, because such a machine needs the ability to generalize some properties beyond what it is forced to be by some given knowledge.

However, in formalizing conjectural reasoning which is needed to generalize properties. the circumscription technique has many merits. First, circumscription is a form of nonmonotonic reasoning. Under the circumstances in which it is always possible to introduce new axioms, any reasoning which jumps to conclusions will be forced to have a nonmonotonic property as far as it is based on logic. Second, circumscription is given a modeltheoretic account, in addition to its proof-theoretic formulation. It can give clear comprehension and insight. Third, the circumscription technique involves predicate variables in its formulation. It becomes, of course, a fault of the technique in that it becomes difficult to implement the reasoning by circumscription, but it can set a logical part of the learning mechanism apart from the extra-logical parts which control the reasoning processes for their efficiency or human preference, for example, some heuristics in learning programs. As a result, problems to be solved will be exposed gradually as extralogical parts change to logical parts. Therefore, the circumscription technique seems to be the most hopeful to formalize various conjectural reasoning uniformly. Here, it still remains to clarify the circumscription technique. We use this term as a formalizing technique using predicate substitution so that adding certain instances of its formulation to given axioms never violates consistency.

Predicate circumscription, as Etherington et al. [1] shows, is not sufficient for realizing default reasoning, and the previous version of circumscription must be extended to more general version, named formula circumscription [7]. Using formula circumscription, we circumscribe a wff, while using predicate circumscription, we circumscribe a predicate. Formula circumscription may be sufficient for the purpose, but unfortunately it is too general. First of all, we show this.

Theorem 1. \dagger A is a formula, P is a tuple of predicate symbols, Φ is a tuple of predicates and E is a wff. $\{\Phi\}$ means a substitution of Φ for P (our notation will be stated more precisely in section 6).

Formula circumscription,

$$A[P] \land \forall \Phi.[(A[\Phi] \land [\forall \mathbf{x}.(E(\mathbf{x}))[\Phi] \supset E(\mathbf{x})] \supset [\forall \mathbf{x}.(E(\mathbf{x}))[\Phi] = E(\mathbf{x})]] \qquad \cdots (1)$$

(modifying McCarthy's terminology), subsumes the following axiom.

[†] This theorem was inspired by a private communication with Etherington.

$$A[P] \wedge (A[\Phi] \supset \forall \mathbf{x}.((P(\mathbf{x}) = \Phi(\mathbf{x}))[\Phi] \equiv (P(\mathbf{x}) = \Phi(\mathbf{x})))) \qquad \cdots \qquad (2)$$

What Theorem 1 means is clearer when Φ does not contain any predicates in P. In this case, $(P(\mathbf{x}) = \Phi(\mathbf{x}))[\Phi]$ is $(\Phi(\mathbf{x}) = \Phi(\mathbf{x}))$, that is, it is true. (1) is therefore simplified to

$$A[P] \wedge (A[\Phi] \supset \forall \mathbf{x}.(P(\mathbf{x}) \equiv \Phi(\mathbf{x}))$$
. ... (3)

(3) is too general, because (3) says that any predicates which can be the substitutes for P are equivalent to P.

Example 1. (reviewing McCarthy's blocks world in [6]) In (3), let P be Block and A be

Let us substitute $\Phi(x) = (x=a)$, in which Block does not occur. This gives

$$\forall x. (Block(x) = (x=a)). \qquad \cdots (5)$$

Again, this means that a concept of 'Block' can be defined without being concerned whether a is a block or not.

This paper proposes a more specific formulation using the circumscription technique. This formulation avoids reasoning by the somewhat forcible means seen in Example 1, and is also guaranteed to preserve consistency, which circumscription does not in general. The more specific formulation formalizes various types of conjectural reasoning uniformly.

3. Ascription Schema

In this paper we write t instead of a tuple of finite terms for brevity. For example, a formula, $A(\mathbf{x})$, stands for $A(x_1, \dots, x_j)$ and the quantifier $\forall \mathbf{x}$ stands for $\forall x_1, \dots, \forall x_k$. By a finite set of formulas $\{F_1, F_2, \dots, F_m\}$, we mean a formula $F_1 \cap F_2 \cap \dots \cap F_m$, where $F_i \cap (i=1,\dots,m)$ is a closed formula obtained from F_i by prefixing \forall with respect to all of the free variables in F_i .

By n-ary predicate, we mean an expression, $\lambda \mathbf{x}.(A(\mathbf{x}))$, where \mathbf{x} is a tuple of n variables and $A(\mathbf{x})$ is a formula in which \mathbf{x} occurs free and no other variables occur free. That is, a predicate is obtained from a formula by λ -abstracting all of the free variables in it.

Let K be a tuple of distinct predicate symbols, $K1,\dots,Kn$, and Ψ a tuple of predicates, $\Psi1,\dots,\Psi n$, where Ki and Ψ i have the same arity. $[\Psi/K]$ means a substitution, representing

Definition [Ascription schema].

Let K be a tuple of distinct predicate symbols, and let Γ be a set of formulas of first order logic containing all predicates in K. The ascription of K to Ψ in $\Gamma[K]$ is the schema

$$\Gamma[\lambda \mathbf{x}.(K(\mathbf{x}) \triangle \Psi(\mathbf{x}))] \land \forall \mathbf{x}.(\Psi(\mathbf{x}) = \Psi(\mathbf{x})[\lambda \mathbf{x}.(K(\mathbf{x}) \triangle \Psi(\mathbf{x}))])$$

$$\land \Gamma[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))] \land \forall \mathbf{x}.(\Psi(\mathbf{x}) = \Psi(\mathbf{x})[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))])$$

$$\supset \forall \mathbf{x}.(K(\mathbf{x}) = \Psi(\mathbf{x})). \qquad \dots \qquad (6)$$

Here \triangle and ∇ represent respectively \wedge and \vee , or \vee and \wedge in each corresponding predicate. Ψ is a tuple of predicates which have the same arity as the corresponding predicates in K. We call the formula on the left side of this schema the *ascribable condition*, written $As(\Gamma, K \sim \Psi)$, and especially $\forall \mathbf{x}.(\Psi(\mathbf{x}) \equiv \Psi(\mathbf{x})[\lambda \mathbf{x}.(K(\mathbf{x}) \triangle \Psi(\mathbf{x}))]) \wedge \forall \mathbf{x}.(\Psi(\mathbf{x}) \equiv \Psi(\mathbf{x})[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))])$ is called the *fixed substitute condition*.

 $\Gamma[\cdots, \lambda \mathbf{x} \mathbf{i}.(Ki(\mathbf{x} \mathbf{i}) \wedge \Psi \mathbf{i}(\mathbf{x} \mathbf{i})), \cdots]$ expresses the condition that all the tuples of entities that can be shown to have a certain property, Ki, by reasoning from certain facts, Γ , can also be shown to have a certain property, $\Psi \mathbf{i}. \Gamma[\cdots, \lambda \mathbf{x} \mathbf{i}.(Ki(\mathbf{x} \mathbf{i}) \vee \Psi \mathbf{i}(\mathbf{x} \mathbf{i})), \cdots]$ is, as far as Ki is concerned, equivalent to the result of replacing $\neg Ki$ by $\lambda \mathbf{x} \mathbf{i}.(\neg Ki(\mathbf{x} \mathbf{i}) \wedge \neg \Psi \mathbf{i}(\mathbf{x} \mathbf{i}))$. Namely, $\Gamma[\cdots, \lambda \mathbf{x} \mathbf{i}.(Ki(\mathbf{x} \mathbf{i}) \vee \Psi \mathbf{i}(\mathbf{x} \mathbf{i})), \cdots]$ expresses the condition that all the tuples of entities that can be shown not to have a property, Ki, can also be shown not to have a certain property, $\Psi \mathbf{i}. \nabla \mathbf{x}.(\Psi(\mathbf{x}) = \Psi(\mathbf{x})[\lambda \mathbf{x}.(K(\mathbf{x}) \vee \Psi(\mathbf{x}))])$ express that each extension of $\Psi \mathbf{i}$ cannot be changed by such changes in K. When we can show that $As(\Gamma, K \sim \Psi)$, namely, the conjunction of these formulas, is true, (6) lets us conclude the formula on the right side, namely that Ki is equivalent to $\Psi \mathbf{i}$.

It is not permitted to add $\forall \Psi$ to Γ . We must use ascription in the following way.

When the formula, p, follows from a set of formulas, Γ , by a complete deduction system of first order logic, we write $\Gamma \vdash p$. Let $\Gamma_{h-1}\{K^h \sim \Psi^h\}$ be $\Gamma_{h-1} \cup \{As(\Gamma_{h-1},K^h \sim \Psi^h) \supset \forall \mathbf{x}.(K^h(\mathbf{x}) \equiv \Psi^h(\mathbf{x}))\}$ ($h = 1,2,\cdots$) and Γ_0 be Γ . Let Γ_h be $\Gamma_{h-1}\{K^h \sim \Psi^h\}$, written $\Gamma\{K^1 \sim \Psi^1; \cdots; K^h \sim \Psi^h\}$. If a finite number n exists such that $\Gamma_n \vdash p$, we write $\Gamma \mid \sim \{K^1 \sim \Psi^1; \cdots; K^n \sim \Psi^n\}$ p and usually abbreviate this as $\Gamma \mid \sim p$.

Example 1 (continued)

In (6), let K be Block and Γ be

Let us substitute $\Psi(x) = (x=a)$. Then from

$$\Gamma[\lambda x.(\operatorname{Block}(x) \land (x=a))] = \operatorname{Block}(a) \qquad \cdots \qquad (8)$$

$$\Gamma[\lambda x.(\operatorname{Block}(x) \lor (x=a))] = (x=a) \lor \Gamma \qquad \cdots \qquad (9)$$

$$\forall x.((x=a) = (x=a)[\lambda x.(\operatorname{Block}(x) \land (x=a))]) = \operatorname{true} \qquad \cdots \qquad (10)$$

$$\forall x.((x=a) = (x=a)[\lambda x.(\operatorname{Block}(x) \lor (x=a))]) = \operatorname{true} \qquad \cdots \qquad (11)$$

$$(6) \text{ gives} \qquad \qquad \text{Block}(a) \supset \forall x.(\operatorname{Block}(x) = (x=a)). \qquad \cdots \qquad (12)$$

Moreover, in $\Gamma\{Block \sim \lambda x.(x=a)\}$ we ascribe 'Block' to ' $\lambda x.(x=b)$ '. In a similar way, we obtain

$$\begin{split} &\Gamma\{\text{Block}\sim\lambda x.(x=a);\;\text{Block}\sim\lambda x.(x=b)\}\\ &=(\text{Block}(a)\;\vee\;\text{Block}(b))\\ &\wedge\;\text{Block}(a)\;\supset\;\forall x.(\;\text{Block}(x)\;\equiv\;(x=a)\;)\\ &\wedge\;\text{Block}(a)\;\supset\;\forall x.(\;\text{Block}(x)\;\equiv\;(x=b)\;). &\cdots (13) \end{split}$$

Therefore, this gives at most

$$\Gamma \mid \sim \forall x. (Block(x) = (x=a)) \lor \forall x. (Block(x) = (x=b)).$$
 ... (14)

Going back to (6), let us discuss the fixed substitute condition,

 $\forall \mathbf{x}.(\ \Psi(\mathbf{x}) \equiv (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x})\triangle \Psi(\mathbf{x}))]\) \land \forall \mathbf{x}.(\ \Psi(\mathbf{x}) \equiv (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x})\nabla \Psi(\mathbf{x}))]\)$. This condition guarantees that ascription preserves consistency, while circumscription which has no precondition corresponding to this does not generally preserve it. Details on consistency are considered more generally in a later section.

We define a class of predicates, named well-candidates, which satisfy this condition.

Definition [well-candidate].

Ψ is called a well-candidate of K in Γ iff the fixed substitute condition is satisfied.

That is, if Ψ is restricted to a well-candidate of K then ascription becomes equivalent to

$$\Gamma[\lambda \mathbf{x}.(K(\mathbf{x}) \triangle \Psi(\mathbf{x}))] \wedge \Gamma[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))] \supset \forall \mathbf{x}.(K(\mathbf{x}) \equiv \Psi(\mathbf{x})).$$
 (15)

Definition { well-candidate form 1 }.

Let K be a n-tuple of predicates, K1,...,Kn, where Ki is a predicate symbol. Ψ , n-tuples of predicates Ψ i (corresponding to Ki), is the well-candidate form 1 of K if each Ψ i can be transformed into expressions of the form $\lambda \mathbf{x}.(\mathrm{Ki}(\mathbf{x}) \wedge \mathrm{Gi}(\mathbf{x}) \vee \mathrm{Hi}(\mathbf{x}))$, where Gi and Hi are predicates in which no predicate symbols in K occur.

Note that each of the forms, H, $\lambda \mathbf{x}.(K(\mathbf{x}) \vee H(\mathbf{x}))$ and $\lambda \mathbf{x}.(K(\mathbf{x}) \wedge G(\mathbf{x}))$, is a well-candidate of K. In example 1, ' $\lambda x.(\mathbf{x}=\mathbf{a})$ ' is a well-candidate of 'Block'.

Prop. 1.

Let Ψ be well-candidate form 1 of K, then Ψ is a well-candidate of K in any axiom, namely,

$$\vdash \ \forall \mathbf{x}.(\ \Psi(\mathbf{x}) \equiv (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \triangle \Psi(\mathbf{x}))]\) \ \land \ \forall \mathbf{x}.(\ \Psi(\mathbf{x}) \equiv (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))]\). \ \cdots \ (16)$$

4. Application of Ascription to Various Types of Conjectural Reasoning

In this section, ascription is applied to various types of reasoning. As stated at the beginning of this paper, ascription represents the flexible notion that the interpretation of a certain property, K, will lie between the extremes of the two. First we give these extremes. They will be useful in understanding the flexibility of the properties of ascription. Then we look over other two types of reasoning, analogy and induction. Ascription is also considered as a form of some kinds of analogy and induction. Finally, ascription is applied to default reasoning.

4.1 Reasoning in the Extremes, Circumscription and Inscription

Ascription subsumes predicate circumscription in the case where an instance of circumscription which is obtained by substituting a predicate satisfying the fixed substitute condition for a predicate variable in it is added to given axioms.

We can derive two significant products from ascription. One product corresponds to predicate circumscription, which formalizes conjectural reasoning based on the closed-world assumption. The other corresponds to the one called inscription in this paper (both circumscription and inscription are instances of formula circumscription, but here, in contrast to predicate circumscription, we call inscription an axiom on the opposite side of predicate circumscription), which formalizes conjectural reasoning such that entities have any properties which they are not denied to have.

<u>Prop. 2.</u> Both formulas, (17), (18) are derived from $\Gamma \cup \{As(\Gamma, K \sim \Psi) \supset \forall \mathbf{x}. (K(\mathbf{x}) \equiv \Psi(\mathbf{x}))\}$, namely,

$$\Gamma \mid \sim_{\{K \sim \Psi\}}$$

$$\forall \mathbf{x}.(\Psi(\mathbf{x}) \supset K(\mathbf{x})) \land \Gamma[\Psi] \land \forall \mathbf{x}.(\Psi(\mathbf{x}) = (\Psi(\mathbf{x}))[\Psi]) \supset \forall \mathbf{x}.(K(\mathbf{x}) = \Psi(\mathbf{x})) \qquad \cdots \qquad (17)$$

(17) is an instance of circumscription with the fixed substitute condition and (18) is an instance of inscription with the fixed substitute condition. If we restrict Ψ to well-candidates of K, (17) is also an instance of predicate circumscription.

Example 2. Various interesting examples of circumscription are given in [6]. So an example of inscription is shown here. Let Γ be "If he is human, he is animate. And if he is human, he can think." This may be described as

$$\{ \forall x. (Human(x) \supset Animate(x)), \forall x. (Human(x) \supset CanThink(x)) \}.$$
 (19)

We ascribe 'Human' to ' λx .(Animate(x) \wedge CanThink(x))' which is a well-candidate of 'Human'. Then,

$$\forall x.(Human(x) \supset Animate(x) \land CanThink(x))$$

$$\equiv \Gamma, \qquad \cdots \qquad (20)$$

 $\Gamma[\lambda x.(Animate(x) \land CanThink(x))]$

$$∀x.(Animate(x) ∧ CanThink(x) ⊃ Animate(x))$$

$$∧ ∀x.(Animate(x) ∧ CanThink(x) ⊃ CanThink(x)), (21)$$

so this inscription gives

$$\forall \mathbf{x}. (\text{Human}(\mathbf{x}) \equiv \text{Animate}(\mathbf{x}) \land \text{CanThink}(\mathbf{x})). \qquad \cdots \qquad (22)$$

This result says " Only human is animate and able to think."

4.2. Analogical Inference

Ascription is a form of a certain class of analogy. According to the notion of ascription, analogy is considered as follows. When a resembles b, where a and b are tuples of entities,

we consider \mathbf{a} and \mathbf{b} to have some common property Ψ . Now let \mathbf{a} have some property K relevant to Ψ in that K and Ψ satisfy the ascribable condition. Then we can infer that \mathbf{b} also has the property K. Here, satisfying the ascribable condition implies at least that we do not know the fact that \mathbf{b} does not have the property K.

In most cases of formalization of analogy, the treatment of resemblance is unsatisfactory. Resemblance is regarded as an atomic relation which cannot be explained. We may say "a is like b" and "b is like c", but likeness may be used in different senses. If we infer "a is like c" using a rule like modus ponens, the inference is against our intuition in general. Therefore, to formalize analogy we should not ignore the common properties on the basis of which we consider that "a is like b". Moreover, in the case of analogizing K(b) from the fact $\{\Psi(a), \Psi(b), K(a)\}$, the relation between K and Ψ should not be ignored, and K and Ψ must satisfy some condition. Let us take an example. A man is like a firework in that both have short lives. Yet we can never infer that a firework can love someone like a man. If the condition that whatever we know to be capable of loving someone has a short life is satisfied, then the inference that a firework can love someone like a man may be justified. If a further condition, that whatever we know to be incapable of loving someone has a long life or is immortal, is satisfied, then it may be even more secure. The ascribable condition requires that these two conditions must be satisfied.

Example 3. Let Γ be "Hector is animate and would be sad if he were burnt, and if Brutus were burnt, he would be sad, too." Namely,

Clearly, $\Gamma \vdash As(\Gamma,Animate \sim \lambda x.(Burnt(x) \supset Sad(x)))$, therefore

$$\forall x.((Burnt(x) \supset Sad(x)) = Animate(x)).$$
 (24)

This says that whoever is sad when burnt is animate. Therefore,

$$\Gamma$$
 [~ Animate(brutus). (25)

Namely, the reasoning, "If Hector and Brutus are burnt then both are sad, and to this extent Hector and Brutus are like each other. Now, Hector is animate so Brutus may also be so", is then a kind of analogy.

4.3. Inductive Inference

Readers may have already noticed that in a theory with the ascription schema it is possible to reason inductively.

Example 4. Let I consist of some instances.

Then $\Gamma \vdash As(\Gamma, Cold \sim \lambda x. Ruddy-faced(matsumoto-san, x))$, therefore

$$\Gamma \mid \sim \forall x. (Cold(x) \equiv Ruddy-faced(matsumoto-san, x)).$$
 (27)

This means that if the system knows it is cold, then it guesses Matsumoto-san will be ruddy-faced, and if he is ruddy-faced, then it expects a cold day. Moreover, if we add the new predicate 'all' which expresses the property of the whole domain, as proposed by McCarthy [6], and let the new extended theory be $\Gamma' \equiv \Gamma \cup \{\forall x.all(x)\}$, then $\Gamma' \vdash As(\Gamma,all \sim \lambda x.Ruddy-faced(matsumoto-san,x))$. So

$$\Gamma' \vdash \forall x. (all(x) = Ruddy-faced(matsumoto-san, x)), \cdots (28)$$

and therefore,

$$\Gamma' \mid \sim \forall x. (Ruddy-faced(matsumoto-san,x)).$$
 (29)

This means that if the system does not know of a day when Matsumoto-san was not ruddy-faced, then it may guess that he is always ruddy-faced.

4.4 Default Reasoning

Let us apply ascription to default reasoning.

Example 5. McCarthy proposed a predicate 'ab' [7], meaning abnormality, to handle common sense reasoning. Here, 'Abn' is used in a similar sense. Let Γ be as follows. We be now whether a bird Tweety can fly or not.

```
\Gamma \equiv \{ \ \forall x. (Bird(x) \land \neg Ab_1(x) \supset Fly(x)), \\ \forall x. (Penguin(x) \supset Ab_1(x)), \\ \forall x. (Penguin(x) \land \neg Ab_2(x) \supset \neg Fly(x)), \\ Bird(Tweety) \} \qquad \cdots \qquad (30)
```

Now we assume that Tweety is as normal as possible, that is we try to minimize its abnormality. Minimizing abnormality corresponds to supplementing lack of knowledge with common sense knowledge. We choose the following candidates corresponding to Ab₁, Ab₂, Fly and Penguin.

$$\begin{array}{l} \prod \lambda x. (Ab_1(x) \wedge (false)), \\ \lambda x. (Ab_2(x) \wedge (false)), \\ \lambda x. (Penguin(x) \wedge (false)), \\ \lambda x. (Fly(x) \vee Bird(x))] \\ \\ \\ \\ \bigcap \lambda x. (Ab_1(x) \vee (false)), \\ \lambda x. (Ab_2(x) \vee (false)), \\ \lambda x. (Penguin(x) \vee (false)), \\ \lambda x. (Fly(x) \wedge Bird(x))] \\ \\ \\ \\ \\ \\ \bigvee \forall x. \neg Ab_1(x) \wedge \forall x. \neg Ab_2(x) \wedge \forall x. \neg Penguin(x) \\ \\ \\ \wedge \forall x. (Fly(x) \equiv Bird(x)) \\ \end{array}$$

$$\begin{array}{c} (31)$$

Therefore

$$\Gamma \mid \sim \forall x.(\text{Fly}(x) = \text{Bird}(x)), \qquad \cdots$$
 (32)

and

$$\Gamma \sim \text{Fly(p-suke)}$$
. (33)

5. Model Theory of Ascription

For the model theory of ascription, we introduce the most \Psi-tending model.

Here, we write |M| for the domain of a model M and M[|P|] for the extension of a predicate P in M.

Definition [more Ψ-tending model in K].

We say M is a more Ψ -tending model of Γ than N in K, writing $M \ge_{K \sim \Psi} N$, if both M and N are Ψ -tending models in K, and

1) |M| = |N|.

200

- 2) M[IPI]=N[IPI] for every predicate symbol P not in K and
- M[|λx.(Ki(x)∧Ψi(x))|] ⊇ N[|λx.(Ki(x)∧Ψi(x))|] and
 M[|λx.(Ki(x)√Ψi(x))|] ⊆ N[|λx.(Ki(x)√Ψi(x))|] for every Ki and Ψi in K and Ψ.

Definition [most Ψ-tending model in K].

A model M of Γ is called most Ψ -tending in K iff $M' \ge_{K \sim \Psi} M$ only if M' = M.

Theorem 2. Any instance of ascription is true in all the most \Psi-tending models in K.

6. Consistency of Ascription

This section shows that any instance of ascription preserves consistency. That is, this guarantees that a consistent Γ cannot contradict the result derived by ascription. Before going any further, we must consider substitution for the predicate. We start by introducing the concept of free substitution.

By a free substitution we mean a substitution which is free in the sense of [2]. A free substitution must satisfy the following two conditions. (A1) In replacing predicate symbols K in a formula, F, by some predicate, Ψ , any variable in each terms, \mathbf{u} , attached to K in F must be free in $\Psi(\mathbf{u})$; and (A2) any free variable (not bound by quantifiers and not λ -abstracted) in Ψ must remain free in $F[\Psi/K]$. However, note that in this paper we regard any predicate as a closed λ -expression, so (A2) is always satisfied. Moreover, by renaming adequate variables in Ψ , we can always ensure that (A1) is satisfied. Therefore, in this paper, a substitution represented by a pair of brackets [...] is always considered to be free.

If S is a set of substitutions, S* denotes the set of all the substitutions which are represented by finite sequence of elements of S.

<u>Theorem 3.</u> Let Θ be a set of free substitution, $\Theta = \{\theta_1, \dots, \theta_n\}, \ 0 \in \Theta^*$ and p be a formula. Then

 $\Gamma \cup \{ \Gamma \theta_1 \wedge \cdots \wedge \Gamma \theta_n \wedge p\theta \supset p \}$ is consistent.

Theorem 4.

- 1) Any instance of ascription of K to Ψ preserves consistency and
- any instance of ascription is true in all the most Ψ-tending models in K.

By <u>Theorem 4</u>, if the antecedent of ascription schema is satisfied, then we are assured of the existence of most Ψ -tending models of K. Note that the result of <u>Theorem 3</u> can be applied to circumscription, and in circumscription, we can similarly think of the well-candidate form which satisfies its fixed substitute condition, that is $\forall \mathbf{x}. (\ \Psi(\mathbf{x}) \equiv (\Psi(\mathbf{x}))[\lambda \mathbf{x}.\Psi(\mathbf{x})]$). Lifschitz showed that circumscription preserves consistency when Γ is a set of almost universal formulas [4], which is a generalized class of separable formulas he proposed himself [3] and universal formulas proposed by Etherington [1]. Note that this condition governs Γ , while the fixed substitute condition governs the predicates which ascription relativizes. However, the couples of predicates intended in [3] to be relativized by circumscription under the separability condition satisfy its fixed substitute condition, because the separability condition requires that no predicate in Ψ may contain predicate in K. From this standpoint, the fixed substitute condition is a weaker condition than the

separability condition. When the fixed substitute condition is satisfied, if the antecedent of circumscription schema is satisfied, even with no minimal model, a most Ψ-tending model exists and circumscription preserves consistency.

6.Conclusion and remarks

As described above, ascription uniformly formalizes diverse and flexible conjectural reasoning. But, of course, there still remain more difficult problems on its use. How do we, humans, use these various types of reasoning properly? Our conclusions will often contradict each other depending on how we interpret our knowledge about a certain property K; in a narrow sense, as in circumscription, or in a broad sense, as in analogy. This problem is deeply relevant to human preference and lies beyond the scope of our logic. We have not considered this much, but it seems that when we have less instances of K, we prefer a narrow interpretation, and that when we have sufficient instances of K, we prefer a broad interpretation. Considered from the viewpoint of ascription, this seems to correspond more or less to the situation that there are, roughly speaking, so many various dubious candidates for Ψ to K in the former case. Indeed, it will be difficult to choose an adequate W, but Kmin is one of the well-founded candidates. In the latter case, because we get more information on K, there are fewer candidates so it seems to be easier to choose. Anyway, an adequate Ψ will usually be given in a moderate sense, i.e., neither in the narrowest nor in the broadest sense. We believe that ascription is a general form which can cover any proper interpretation of K between one extreme and another.

ACKNOWLEDGMENTS

I would like to thank Dr. Koichi Furukawa especially, and the other members of the First Research Laboratory and AAL Working Group at ICOT for their useful comments. Also, I wish to express my gratitude to Dr. Kazuhiro Fuchi, Director of the ICOT Research Center, who provided me with the opportunity to pursue this research.

REFERENCES

- [1] Etherington, D., Mercer, R. and Reiter, R.: On the adequacy of predicate circumscription for closed-world reasoning, Technical report 84-5, Dept. of Computer Science, Univ. of British Columbia (1984).
- [2] Kleene, S.C.: Introduction to Metamathematics, North-Holland, 1971, CH. WI.
- [3] Lifschitz, V.: Computing circumscription, in: Proceedings of Ninth International Joint Conference on Artificial Intelligence, Los Angeles, CA (1985) 121-127.
- [4] Lifschitz, V.: On the Satisfiability of Circumscription, Artificial Intelligence 28 (1986) 17-27.
- [5] Michalski, R.S., Carbonell, J.G. and Mitchell, T.M.: Machine Learning An Artificial Intelligence Approach, Tioga, 1983, chap. 4.
- [6] McCarthy, J.: Circumscription a form of non-monotonic reasoning, Artificial Intelligence 13 (1980) 27-39.
- [7] McCarthy, J.: Application of circumscription to formalizing common-sense knowledge, Artificial Intelligence 28 (1986) 89-116.
- [8] McDermott, D. and Doyle, J.: Non-monotonic Logic I, Artificial Intelligence 13 (1980) 41-72.

APPENDIX

Some of the proofs in this paper use Kleene's theorems [2]. The following theorems, K1 and K2, are obtained directly from Kleene's theorems.

Theorem K1. A(x), B(x) and C(x) are formulas.

$$\forall \mathbf{x}.(\mathbf{A}(\mathbf{x}) \equiv \mathbf{B}(\mathbf{x})) \vdash \forall \mathbf{x}.(\mathbf{C}(\mathbf{x})[\mathbf{A}/\mathbf{P}] \equiv \mathbf{C}(\mathbf{x})[\mathbf{B}/\mathbf{P}]).$$

Theorem K2. Let Γ be a set of closed formulas, E be a closed formula and θ be a free (predicate) substitution.

If
$$\Gamma \vdash E$$
 then $\Gamma\theta \vdash E\theta$.

<u>Theorem 1.</u> A is a formula, P is a tuple of predicate symbols, Φ is a tuple of predicates and E is a closed formula.

Formula circumscription,

$$A[P] \land \forall \Phi . ((A[\Phi] \land (\forall \mathbf{x} . (E(\mathbf{x}))[\Phi] \supset E(\mathbf{x})) \supset (\forall \mathbf{x} . (E(\mathbf{x}))[\Phi] = E(\mathbf{x}))), \cdots (1)$$

subsumes the following axiom.

$$A[P] \land (A[\Phi] \supset \forall \mathbf{x}.((P(\mathbf{x}) = \Phi(\mathbf{x}))[\Phi] = (P(\mathbf{x}) = \Phi(\mathbf{x})))) \qquad \cdots \qquad (2)$$

<u>Proof.</u> First, consider a part of (1), the formula $(\forall \mathbf{x}.(E(\mathbf{x}))[\Phi] \supset E(\mathbf{x}))$. Let $E(\mathbf{x})$ be $\neg (P(\mathbf{x}) \equiv \Phi(\mathbf{x}))$ for some Φ . Then $\forall \mathbf{x}.(\neg (P(\mathbf{x}) \equiv \Phi(\mathbf{x})))[\Phi] \supset \neg (P(\mathbf{x}) \equiv \Phi(\mathbf{x}))$. Namely $\forall \mathbf{x}.((P(\mathbf{x}) \equiv \Phi(\mathbf{x})) \supset (\Phi(\mathbf{x}) \equiv \Phi(\mathbf{x})[\Phi]))$. And $\forall \mathbf{x}.((P(\mathbf{x}) \equiv \Phi(\mathbf{x})) \supset (\Phi(\mathbf{x})[P] \equiv \Phi(\mathbf{x})[\Phi]))$. This is valid by theorem K1. So from (1) we can obtain

$$A[P] \land (A[\Phi] \supset \forall \mathbf{x}.((P(\mathbf{x}) \equiv \Phi(\mathbf{x}))[\Phi] \equiv (P(\mathbf{x}) \equiv \Phi(\mathbf{x}))).$$

Prop. 1.

Let Ψ be well-candidate form 1 of K, then Ψ is a well-candidate of K in any axioms, namely

```
\vdash \forall \mathbf{x}.(\ \Psi(\mathbf{x}) = (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(\mathbf{K}(\mathbf{x}) \triangle \Psi(\mathbf{x}))]\ ) \land \forall \mathbf{x}.(\ \Psi(\mathbf{x}) = (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(\mathbf{K}(\mathbf{x}) \nabla \Psi(\mathbf{x}))]\ ). \cdots (16)
```

Proof. By predicate calculus.

<u>Prop. 2.</u> Both formulas, (17), (18) are derived from $\Gamma \cup \{As(\Gamma, K \sim \Psi) \supset \forall \mathbf{x}. (K(\mathbf{x}) \equiv \Psi(\mathbf{x}))\}$, namely,

 $\Gamma \mid \sim_{\{K \sim \Psi\}}$

$$\forall \mathbf{x}.(\Psi(\mathbf{x}) \supset K(\mathbf{x})) \land \Gamma[\Psi] \land \forall \mathbf{x}.(\Psi(\mathbf{x}) = (\Psi(\mathbf{x}))[\Psi]) \supset \forall \mathbf{x}.(K(\mathbf{x}) = \Psi(\mathbf{x})) \qquad \cdots \qquad (17)$$

<u>Proof</u>. By predicate calculus.

Theorem 2. Any instance of ascription is true in all the most Ψ-tending models in K.

Proof. Let M be a most Ψ -tending models of Γ in K. Now let the left side of (1) be satisfied. If the right side of (1) were not satisfied, the extension of Ki would not be the same to ones of $\lambda x.(Ki(x) \wedge \Psi i(x))$ or $\lambda x.(Ki(x) \vee \Psi i(x))$. Then let the extension of Ki not be the same to ones of $\lambda x.(Ki(x) \wedge \Psi i(x))$. In that case, we could get a more Ψ -tending model N (\neq M) such that N[|Ki|] = M[|\lambda x.(Ki(x) \lambda \Psi i(x))|] and N[|\Psi i|] = M[|\Psi i|]. Because

 $M[|\lambda x.(Ki(x) \wedge \Psi i(x))|] = M[|\lambda x.(Ki(x) \wedge \Psi i(x))|] \cap M[|\Psi i|]$

- = $N[|Ki|] \cap N[|\Psi i|]$
- = $N[|\lambda x.(Ki(x) \wedge \Psi i(x))|]$

 $M[|\lambda x.(Ki(x) \vee \Psi i(x))|] \supset M[|\Psi i|]$

- = $M[!\lambda x.(Ki(x) \wedge \Psi i(x))] \cup M[!\Psi i!]$
- = N[tKit] ∪ N[tΨit]
- = $N[|\lambda x.(Ki(x) \vee \Psi i(x))|]$,

therefore $N >_{K \sim \Psi} M$. This contradicts the assumption that M is a most Ψ -tending models of Γ in K.

<u>Prop. A1.</u> Let Γ be a set of formulas of first order logic and ρ be a closed formula of first order logic. If Γ is consistent and some free substitution θ exists such that $\Gamma \vdash \Gamma\theta \land \rho\theta$, then $\Gamma \cup \{\rho\}$ is consistent.

<u>Proof.</u> Let θ be a free substitution such that $\Gamma \vdash \Gamma \theta \land p\theta$. Now we assume that $\Gamma \cup \{p\} \vdash \Box$ (representing 'false'). Namely $\Gamma \vdash \neg p$. Using theorem K2, $\Gamma \theta \vdash \neg p\theta$ follows. Here $\Gamma \vdash \Gamma \theta$, so $\Gamma \vdash \neg p\theta$. This result contradicts the assumption that $\Gamma \vdash p\theta$.

Theorem 3. Let Θ be a set of free substitution, $\{\theta_1, \dots, \theta_n\}$, $\theta \in \Theta^*$ and p be a closed formula.

Then $\Gamma \cup \{ \Gamma\theta_1 \wedge \cdots \wedge \Gamma\theta_n \wedge p\theta \supset p \}$ is consistent.

<u>Proof.</u> Assume that $\Gamma \vdash \Gamma\theta_1 \land \cdots \land \Gamma\theta_n \land p\theta \land \neg p$. Using theorem K2 repeatedly, we can obtain $\Gamma \vdash \Gamma\theta$. Therefore from **prop.** A1, $\Gamma \cup \{p\}$ is consistent. This is contradicts the assumption that $\Gamma \vdash \neg p$.

```
Prop. A2.
```

```
If \Gamma \vdash \forall \mathbf{x}. (\Psi(\mathbf{x}) = (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \land \Psi(\mathbf{x}))] ) \land \forall \mathbf{x}. (\Psi(\mathbf{x}) = (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))] ) then \Gamma \vdash \forall \mathbf{x}.(K(\mathbf{x}) = \Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \Delta \Psi(\mathbf{x}))] \cdot [\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))].

Proof. Let [\lambda \mathbf{x}.(K(\mathbf{x}) \Delta \Psi(\mathbf{x}))] be \{\cdots, \lambda \mathbf{x}.(Ki(\mathbf{x}) \land \Psi i(\mathbf{x})), \cdots].

\vdash \forall \mathbf{x}.(K(\mathbf{x}) = \Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \Delta \Psi(\mathbf{x}))] = \land_i \forall \mathbf{x}.((Ki(\mathbf{x}) \land \Psi i(\mathbf{x})) = (\Psi i(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \Delta \Psi(\mathbf{x}))])

From \Gamma \vdash \forall \mathbf{x}.(\Psi(\mathbf{x}) = (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \Delta \Psi(\mathbf{x}))]),

\Gamma \vdash \forall \mathbf{x}.(K(\mathbf{x}) = \Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \Delta \Psi(\mathbf{x}))] = \land_i \forall \mathbf{x}.((Ki(\mathbf{x}) \land \Psi i(\mathbf{x})) = \Psi i(\mathbf{x})).

Now using theorem K1 and \Gamma \vdash \forall \mathbf{x}.(\Psi(\mathbf{x}) = (\Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))]),

\Gamma \vdash \forall \mathbf{x}.(K(\mathbf{x}) = \Psi(\mathbf{x}))[\lambda \mathbf{x}.(K(\mathbf{x}) \Delta \Psi(\mathbf{x}))] \cdot [\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))]

= \land_i \forall \mathbf{x}.((Ki(\mathbf{x}) \land \Psi i(\mathbf{x})) \Rightarrow \Psi i(\mathbf{x})[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))]) = \Psi i(\mathbf{x})[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))])

= \land_i \forall \mathbf{x}.((Ki(\mathbf{x}) \lor \Psi i(\mathbf{x})) \land \Psi i(\mathbf{x})[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))]) = \Psi i(\mathbf{x})[\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))])

= \land_i \forall \mathbf{x}.((Ki(\mathbf{x}) \lor \Psi i(\mathbf{x})) \land \Psi i(\mathbf{x})) \Rightarrow \Psi i(\mathbf{x}))

= \land_i \forall \mathbf{x}.((Ki(\mathbf{x}) \lor \Psi i(\mathbf{x})) \land \Psi i(\mathbf{x})) \Rightarrow \Psi i(\mathbf{x}))

= \land_i \forall \mathbf{x}.((Ki(\mathbf{x}) \lor \Psi i(\mathbf{x})) \land \Psi i(\mathbf{x})) \Rightarrow \Psi i(\mathbf{x}))

= \land_i \forall \mathbf{x}.((Ki(\mathbf{x}) \lor \Psi i(\mathbf{x})) \land \Psi i(\mathbf{x})) \Rightarrow \Psi i(\mathbf{x}))
```

Theorem 4.

- 1) Any instance of ascription of K to Ψ preserves consistency and
- 2) any instance of ascription is true in all the most Ψ-tending models in K.

Proof. In Theorem 3 let $\Theta = \{\theta_1, \theta_2\}, \ \theta_1 = [\lambda \mathbf{x}.(K(\mathbf{x}) \triangle \Psi(\mathbf{x}))], \ \theta_2 = [\lambda \mathbf{x}.(K(\mathbf{x}) \nabla \Psi(\mathbf{x}))]$ and $\mathbf{p} = \forall \mathbf{x}.(K(\mathbf{x}) = \Psi(\mathbf{x})).$ From Prop. A2, $\Gamma \vdash \forall \mathbf{x}.(\Psi(\mathbf{x}) = (\Psi(\mathbf{x}))\theta_1) \land \forall \mathbf{x}.(\Psi(\mathbf{x}) = (\Psi(\mathbf{x}))\theta_2) \land \forall \mathbf{x}.(\Psi(\mathbf{x}) = \Psi(\mathbf{x})) \land \forall \mathbf{x}.(\Psi(\mathbf{x}) = (\Psi(\mathbf{x}))\theta_2) \land \forall \mathbf{x}.(\Psi(\mathbf{x}) = \Psi(\mathbf{x})) \land \forall$