ICOT Technical Report: TR-440

TR-440

Formalizing Nonmonotonic Reasoning by Preference Order

by

K. Sato

December, 1988

ICOT

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

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

Formalizing Nonmonotonic Reasoning by Preference Order

Ken Satoh

ICOT
1-4-28, Mita, Minato-ku, Tokyo 108, Japan

csnet: ksatoh@icot.jp
uucp: {enea,inria,kddlab,mit-eddie,ukc}!icot!ksatoh

December 6, 1988

ABSTRACT

Recently, some researchers found that the previous formalisms of *nonmonotonic reasoning* such as predicate circumscription or default logic are not enough to capture some commonsense reasoning. We believe that the cause of the problems is because those formalism cannot represent the preference order in those problems which human possesses.

In this paper, we give a formalism of nonmonotonic reasoning by defining a meta-language to represent preference order over interpretations of second-order language and translating it to the second-order language to provide a proof theory. By this formalism, we can express broader classes of preferences over interpretations and infer results syntactically by the second-order sentences of its translation.

We first define the model-theoretic meta-language to express relations over interpretations and show how to translate expression in meta-language into second-order language. Then we show some examples of formalization of nonmonotonic reasoning. Finally, we discuss the limitations of our framework in two aspects, that is, reasoning of inequality and conditional probability.

Keywords: Nonmonotonic Reasoning, Preference Order, Circumscription.

1 Introduction

In real life, we are sometimes forced to make a decision even if there is not enough information. For one solution to those situations, we use our commonsense to complement unknown information. Commonsense is not logically true, but practically, it works well because commonsense is a collection of normal results. However, since commonsense is not logically true, it may be false when more information is obtained. In this case, the result derived from commonsense must be invalidated. Such reasoning is called *nonmonotonic reasoning* because the derived result does not increase monotonically as more information is obtained. This phenomenon has been formalized by various researchers [McCarthy80, McDermott80, Reiter80].

Unfortunately, those formalism does not capture a nonmonotonic reasoning in inheritance system and temporal projection. [McCarthy86] points out that a simple abnormality formalism does not work in simple circumscription and introduces prioritized circumscription. [Etherington86] also points out that the normal default theory does not produce a unique extension in the inheritance system and introduces the semi-normal default theory. [Hanks87] points out that simple circumscription, normal default theory and NML-I do not capture a nonmonotonic reasoning in the *Yale shooting problem*. We believe that the reason why those problems emerge in the previous formalisms is that they can not express the preferences in those problems whereas human can express them. Several researchers suggest solutions to the Yale shooting problem along this line [Kautz86, Lifschitz86, Shoham86]. We also give a proposal which not only solves the Yale shooting problem but also a problem in the inheritance system by translating the problems into reasoning in tree-structured multiple worlds and identifying preference order of those problems as a order where a preferred model changes minimally in one direction [Satoh87] [1].

In this paper, we generalize our solution so that we can express broader classes of preferences over interpretations. We define a model-theoretic meta-language to describe a relation over interpretations of second-order predicate logic. Then, we show how to translate the relation described in the meta-language into the second-order sentences. If we apply this framework to the preference order over models of a closed second-order sentence, the order is shown to be translated into a generalized form of circumscription[Lifschitz84]. Therefore, this paper can be regarded as giving a formal semantics for general circumscription.

Recently, Shoham[Shoham87] also gave a semantical framework on various formalisms of nonmonotonic reasoning including his solution to the Yale shooting problem. His framework is to define a new logic by augmenting a standard logic by introducing a preference over its interpretations. Although his framework is very general, his definition is only semantical one and there is no proof theory. Therefore, this paper also can be regarded as giving a proof theory for his framework in the second-order logic.

[1] We shall explain this solution briefly in Section 5.5.

The structure of the paper is as follows. The first three sections are related to the formal definition of the model-theoretic meta-language and the translation of the meta-language into the second-order sentences. Then, we show how to apply this framework to formalizing nonmonotonic reasoning by using some examples. Finally, we discuss the limitation of this framework.

2 Definition of the language and interpretations

In the second-order logic, we shall use commas, parentheses, the symbols of the logical connectives $\neg$, $\supset$, $\wedge$, $\vee$ and $\equiv$, the quantifier symbols $\forall$ and $\exists$, and the following groups of symbols.

Individual constants: $a, b, c, \ldots.$
Individual variables: $x, y, z, \ldots.$
Function constants: $f, g, h, \ldots.$
Predicate constants: $P, Q, R, \ldots.$
Predicate variables: $p, q, r, \ldots.$

When we say a variable v, v is an individual variable or a predicate variable.

Definition 1 *Terms*

1. Individual variables and individual constants are terms.
2. If f^n is a n-ary function constant and $t_1, \ldots, t_n$ are terms, then $f^n(t_1, \ldots, t_n)$ is a term.
3. An expression is a term only if it satisfies one of the above conditions.

Definition 2 *Well-formed formulas (wffs)*

1. If P^n is a n-ary predicate constant and $t_1, \ldots, t_n$ are terms, then $P^n(t_1, \ldots, t_n)$ is a wff.
2. If p^n is a n-ary predicate variable and $t_1, \ldots, t_n$ are terms, then $p^n(t_1, \ldots, t_n)$ is a wff.
3. If A and B are wffs and v is a variable, then $\neg A$, $A \supset B$ and $\forall v A$ are wffs.
4. An expression is a wff only if it satisfies one of the above conditions.

If A and B are wffs and v is a variable, then $A \wedge B$, $A \vee B$, $A \equiv B$ and $\exists v A$ are abbreviations for $\neg(A \supset \neg B)$, $\neg A \supset B$, $(A \supset B) \wedge (B \supset A)$ and $\neg(\forall v \neg A)$, respectively. We call a wff without any predicate variables a *first-order wff*. And we call a wff without any free variables a *closed wff*.

We write a wff A with some of the free variables $v_1, \ldots, v_n$ as $A(v_1, \ldots, v_n)$. Then we write as $A(t_1, \ldots, t_n)$ the result of substituting in A the terms $t_1, \ldots, t_n$ for all free occurrences of $v_1, \ldots, v_n$, respectively.

An *interpretation* M consists of a nonempty set D, called the *domain* of the interpretation, and an assignment to each individual constant a of an element $(a)^M$ of D, to each n-ary function constant f^n of a function $(f^n)^M$ from D^n to D, and to each n-ary predicate constant P^n of a subset $(P^n)^M$ of D^n. Individual variables range over the set D and n-ary predicate variables vary over the powerset of D^n.

We consider an assignment function ϕ (with respect to the domain D) from an individual variable and a n-ary predicate variable to an element of D and a subset of D^n, respectively. We denote all assignment functions (with respect D) as Φ_D. We denote an assignment function ϕ_v which differs from ϕ at most the assignment of the variable v. We write as $\phi_{v_1...v_n}$ which differs from $\phi_{v_1...v_{n-1}}$ at most the assignment of v_n.

Let t be a term. We extend an assignment function ϕ to a function ϕ^M that assigns to each term t an element $\phi^M(t)$ in D as follows.

1. If t is an individual constant, then $\phi^M(t) = (t)^M$
2. If t is an individual variable, then $\phi^M(t) = \phi(t)$.
3. If t is of the form $f^n(t_1, ..., t_n)$, then $\phi^M(t) = (f^n)^M(\phi^M(t_1), ..., \phi^M(t_n))$.

Let M be an interpretation with domain D. An assignment function ϕ (with respect to the domain D) satisfies a wff A in M (written as $M \models_\phi A$) if and only if the following conditions are satisfied.

1. If A is of the form $P^n(t_1, ..., t_n)$ where P^n is a n-ary predicate constant, then $< \phi^M(t_1), ..., \phi^M(t_n) > \in (P^n)^M$.
2. If A is of the form $p^n(t_1, ..., t_n)$ where p^n is a n-ary predicate variable, then $< \phi^M(t_1), ..., \phi^M(t_n) > \in \phi(p^n)$.
3. If A is of the form $\neg B$, not $M \models_\phi B$ (written as $M \not\models_\phi B$).
4. If A is of the form $B \supset C$, either $M \not\models_\phi B$ or $M \models_\phi C$.
5. If A is of the form $\forall v B$ where v is a variable, for every ϕ_v in Φ_D, $M \models_{\phi_v} B$.

Let M be an interpretation with the domain D. A wff A is true for the interpretation M (written as $M \models A$) if and only if for every assignment function ϕ in Φ_D, $M \models_\phi A$. A wff A is false for the interpretation M if and only if for every assignment function ϕ in Φ_D, $M \not\models_\phi A$. An interpretation M is said to be a *model* for a set Γ of wffs if and only if every wff in Γ is true for M.

Proposition 1 *Let M be an interpretation with the domain D. If every variable in a term t is one of $x_1, ..., x_n$, and if ϕ and ϕ' are the assignment functions with respect to D such that for every x_i, $\phi(x_i) = \phi'(x_i)$, then $\phi^M(t) = \phi'^M(t)$.*

Proof. We prove the above by induction on the number m of function letter in t. Assume the result holds for all integers $< m$.

1. If t is an individual constant a, then $\phi^M(a) = (a)^M = \phi'^M(a)$.

2. If t is an individual variable x, then $\phi^M(x) = \phi(x) = \phi'(x) = \phi'^M(x)$.

3. If t is of the form $f^n(t_1, ..., t_n)$. Each t_i has fewer than m function letters. By inductive hypothesis, $\phi^M(t_i) = \phi'^M(t_i)$. Then $\phi^M(f^n(t_1, ..., t_n)) = (f^n)^M(\phi^M(t_1), ..., \phi^M(t_n)) = (f^n)^M(\phi'^M(t_1), ..., \phi^M(t_n)) = \phi'^M(f^n(t_1, ..., t_n))$. □

Proposition 2 *Let A be a wff all of whose free variables are $v_1, ..., v_n$ and M be an interpretation and ϕ and ϕ' be assignment functions. If for every v_i, $\phi(v_i) = \phi'(v_i)$, then $M \models_\phi A$ if and only if $M \models_{\phi'} A$.*

Proof. We prove the above by induction on the number r of connectives and quantifiers in A. Assume the result holds for all integers $< r$.

1. A is of the form $P^n(t_1, ..., t_n)$ where P^n is a n-ary predicate constants. By the proposition 1, for each t_i, $\phi^M(t_i) = \phi'^M(t_i)$. Therefore $M \models_\phi P^n(t_1, ..., t_n)$ if and only if $< \phi^M(t_1), ..., \phi^M(t_n) > \in (P^n)^M$ if and only if $< \phi'^M(t_1), ..., \phi'^M(t_n) > \in (P^n)^M$ if and only if $M \models_{\phi'} P^n(t_1, ..., t_n)$

2. A is of the form $p^n(t_1, ..., t_n)$ where p^n is a n-ary predicate variables. By the condition of ϕ and ϕ', $\phi(p^n) = \phi'(p^n)$. By the proposition 1, for each t_i, $\phi^M(t_i) = \phi'^M(t_i)$. Therefore $M \models_\phi p^n(t_1, ..., t_n)$ if and only if $< \phi^M(t_1), ..., \phi^M(t_n) > \in \phi(p^n)$ if and only if $< \phi'^M(t_1), ..., \phi'^M(t_n) > \in \phi'(p^n)$ if and only if $M \models_{\phi'} p^n(t_1, ..., t_n)$

3. A is of the form $\neg B$. B has fewer than r connectives and quantifications and B does not contain any predicate constant in $\mathbf{P}$. By the inductive hypothesis, $M \models_\phi B$ if and only if $M \models_{\phi'} B$. By the definition of the satisfaction, $M \models_\phi \neg B$ if and only if $M \not\models_\phi B$ and $M \models_{\phi'} \neg B$ if and only if $M \not\models_{\phi'} B$. Therefore $M \models_\phi \neg B$ if and only if $M \models_{\phi'} \neg B$.

4. A is of the form $B \supset C$. This case is proved in a similar way to the previous case.

5. A is of the form $\forall v B$ where v is a variable. Assume $M \models_\phi A$. Then for every ϕ_v in Φ_D, $M \models_{\phi_v} B$. Take any ϕ'_v. Then we can take some ϕ_v such that $\phi_v(v) = \phi'_v(v)$. Then ϕ_v and ϕ'_v agree on the assignments of $v, v_1, ..., v_n$ some of which are free variables in B. By the inductive hypothesis that $M \models_{\phi_v} B$ if and only if $M \models_{\phi'_v} B$. Since for every ϕ_v in Φ_D, $M \models_{\phi_v} B$, for every ϕ'_v in Φ_D, $M \models_{\phi'_v}$. Hence $M \models_{\phi'} A$. The converse also holds in a similar way. □

Note that if A is a closed wff, then $M \models_\phi A$ if and only if $M \models_{\phi'} A$, and hence $M \models A$.

We will later use the following proposition.

Proposition 3 *There exists $\phi_{v_1...v_n}$ in Φ_D such that $M \models_{\phi_{v_1...v_n}} A$ if and only if $M \models_\phi \exists v_1 ... \exists v_n A$.*

Proof. We prove the above by induction on the number n of variables in $\phi_{v_1...v_n}$. Assume the result holds for all integers $< n$.

1. In case of $n = 1$. By the definition of satisfaction, there exists ϕ_{v_1} in Φ_D such that $M \models_{\phi_{v_1}} A$ if and only if $M \models_\phi \exists v_1 A$.

2. In case of $n > 1$. Assume there exists $\phi_{v_1...v_n}$ in Φ_D such that $M \models_{\phi_{v_1...v_n}} A$. Then, by the definition of satisfaction, $M \models_{\phi_{v_1...v_{n-1}}} \exists v_n A$. Therefore, there exists $\phi_{v_1...v_{n-1}}$ in Φ_D such that $M \models_{\phi_{v_1...v_{n-1}}} \exists v_n A$. By the inductive hypothesis, $M \models_\phi \exists v_1 ... \exists v_{n-1}(\exists v_n A)$.

   Assume $M \models_\phi \exists v_1 \exists v_2 ... \exists v_n A$. Then, by the definition of satisfaction, There exists ϕ_{v_1} such that $M \models_{\phi_{v_1}} \exists v_2 ... \exists v_n A$. By the inductive hypothesis, There exists $\phi_{v_1...v_n}$ such that $M \models_{\phi_{v_1...v_n}} A$. □

3 The definition of the language for model theoretic statements

In the previous section, we used everyday English to talk about model theoretic statement. Here, we use an abbreviation of English sentences on the model theoretic statement. Then we first define *meta-formula* with respect to an interpretation M with the domain D as follows.

Definition 3 *Meta-formula*

1. If A is a wff and ϕ is an assignment function, $M \models_\phi A$ is a meta-formula.

2. If $\mathcal{A}$ and $\mathcal{B}$ are meta-formulas, ϕ is an assignment function and v is a variable, then $\neg\mathcal{A}$, $\mathcal{A} \supset \mathcal{B}$, $(\forall\phi_v \in \Phi_D)\mathcal{A}$ are meta-formulas.

3. An expression is a meta-formula only if it satisfies one of the above conditions.

A *subformula* of a meta-formula $\mathcal{A}$ is defined as follows.

1. $\mathcal{A}$ itself is a subformula of $\mathcal{A}$.

2. If $\mathcal{A}$ is of the form $\neg\mathcal{B}$, then subformulas of $\mathcal{B}$ is subformulas of $\mathcal{A}$.

3. If $\mathcal{A}$ is of the form $\mathcal{B} \supset \mathcal{C}$, then subformulas of $\mathcal{B}$ and $\mathcal{C}$ are subformulas of $\mathcal{A}$.

4. If $\mathcal{A}$ is of the form $(\forall\phi_v \in \Phi_D)\mathcal{B}$, then subformulas of $\mathcal{B}$ is subformulas of $\mathcal{A}$.

A meta-formula $M \models_\phi A$ is called an *atomic meta-formula* and a meta-formula $(\forall\phi_v \in \Phi_D)\mathcal{A}$ is called a *quantified meta-formula*. $\mathcal{A}$ of a quantified meta-formula $(\forall\phi_v \in \Phi_D)\mathcal{A}$ is called the *scope* of the quantifier "$(\forall\phi_v \in \Phi_D)$". If $\mathcal{B}$ is a subformula of a meta-formula $\mathcal{A}$, then the number of quantification of $\mathcal{B}$ is called the *depth* of $\mathcal{B}$ and is defined as follows.

1. The depth of a formula $\mathcal{A}$ itself is 0.

2. If the depth of subformula $\neg\mathcal{B}$ is m, then the depth of subformula $\mathcal{B}$ is m.

3. If the depth of subformula $\mathcal{B} \supset \mathcal{C}$ is m, then the depth of subformulas $\mathcal{B}$ and $\mathcal{C}$ is m.

4. If the depth of subformula $(\forall\phi_v\in\Phi_D)\mathcal{B}$ is m, then the depth of subformula $\mathcal{B}$ is $m+1$.

If the depth of a quantified meta-formula $(\forall\phi_v\in\Phi_D)\mathcal{B}$ is m, then the *depth of the scope* $\mathcal{B}$ is defined as $m+1$.

If $\mathcal{A}$ and $\mathcal{B}$ are meta-formulas, ϕ is an assignment function and v is a variable, then $\mathcal{A} \wedge \mathcal{B}$, $\mathcal{A} \vee \mathcal{B}$, $\mathcal{A} \equiv \mathcal{B}$ and $(\exists\phi_v\in\Phi_D)\mathcal{A}$". are abbreviations for $\neg(\mathcal{A} \supset \neg\mathcal{B})$, $\neg\mathcal{A} \supset \mathcal{B}$, $(\mathcal{A} \supset \mathcal{B}) \wedge (\mathcal{B} \supset \mathcal{A})$ and $\neg((\forall\phi_v\in\Phi_D)\neg\mathcal{A})$, respectively.

Do not confuse the object logical connectives and the meta logical connectives. The meta logical connectives are the following abbreviations of everyday English words.

1. $M \models_\phi A$ means "ϕ satisfies A in M".

2. $\neg\mathcal{A}$ means "not $\mathcal{A}$".

3. $\mathcal{A} \supset \mathcal{B}$ means "either not $\mathcal{A}$ or $\mathcal{B}$".

4. $(\forall\phi_v\in\Phi_D)\mathcal{A}$ means "for every ϕ_v in Φ_D which differs from ϕ at most the assignment of v, $\mathcal{A}$".

Actually we would like to define the particular meta-formulas which are used in the discussion of satisfaction. We call them *meta-well-formed formulas(m-wff)*.

Definition 4 *Meta-well-formed formulas (m-wff)*

A meta-formula $\mathcal{A}$ is a m-wff if and only if the following conditions are satisfied.

1. Every atomic meta-formula in $\mathcal{A}$ whose depth is 0 has the same assignment function ϕ. We call this ϕ the *toplevel assignment function(taf)*. We write as $\mathcal{A}_\phi$ a m-wff with a taf ϕ.

2. For every quantified meta-formula $(\forall\phi'_v\in\Phi_D)\mathcal{B}$ whose depth is 0, ϕ'_v is different from the taf ϕ at most in the assignment of a variable v, that is $\phi'_v = \phi_v$.

3. In every scope of the quantifier "$(\forall\phi_v\in\Phi_D)$" in $\mathcal{A}$ whose depth is m, every atomic meta-formula in the scope whose depth is m has the same assignment function ϕ_v in the quantifier. We call this ϕ_v the *assignment function of the scope*.

4. In every scope of the quantifier "$(\forall\phi_v\in\Phi_D)$" in $\mathcal{A}$ whose depth is m and whose assignment function is ϕ_v, for every quantified meta-formula in the scope $(\forall\phi'_u\in\Phi_D)\mathcal{B}$ whose depth is m, ϕ'_u is different from ϕ_v at most in the assignment of a variable u, that is, $\phi'_u = \phi_{vu}$.

Note that a subformula $\mathcal{B}$ of a m-wff $\mathcal{A}$ is also a m-wff. We sometimes write as $\mathcal{B}_{\phi'}$ a subformula of a m-wff with a taf ϕ'. (For example, if $\mathcal{A}_\phi$ is of the form $\neg\mathcal{B}$, we also write $\neg\mathcal{B}_\phi$)

Example 1 *Meta-well-formed formulas (m-wff)*

- $(\forall\phi_x \in \Phi_D)(M \models_\phi P(x) \wedge M \models_{\phi_x} Q(x))$ is a meta-formula but not a m-wff because this formula violates condition 2.
- $M \models_\phi P(x) \wedge (\forall\phi_{xy} \in \Phi_D)(M \models_{\phi_{xy}} Q(x))$ is a meta-formula but not a m-wff because this formula violates condition 3.
- $M \models_\phi p(x,y) \wedge (\forall\phi_x \in \Phi_D)(M \models_{\phi_x} Q(x) \wedge (\forall\phi_{xx} \in \Phi_D)(M \models_{\phi_{xx}} (q(x) \wedge R(x,z))))$ is a m-wff.

By the definition of m-wff, we can construct an atomic m-wff from any m-wff. Consider the following translation.

Translation 1: from a m-wff to an atomic m-wff
Let $\mathcal{A}_\phi$ be a m-wff with the taf ϕ.

1. $\mathcal{A}_\phi$ is of the form $M \models_\phi A$. It is translated into itself.
2. $\mathcal{A}_\phi$ is of the form $\neg\mathcal{B}_\phi$. It is translated into $M \models_\phi \neg B$, where $\mathcal{B}_\phi$ is translated into $M \models_\phi B$.
3. $\mathcal{A}_\phi$ is of the form $\mathcal{B}_\phi \supset \mathcal{C}_\phi$. It is translated into $M \models_\phi B \supset C$, where $\mathcal{B}_\phi$ is translated into $M \models_\phi B$ and $\mathcal{C}_\phi$ is translated into $M \models_\phi C$.
4. $\mathcal{A}_\phi$ is of the form $(\forall\phi_v \in \Phi_D)\mathcal{B}_{\phi_v}$. It is translated into $M \models_\phi \forall v B$, where $\mathcal{B}_{\phi_v}$ is translated into $M \models_{\phi_v} B$.

Example 2 *Translation from a m-wff into an atomic m-wff*

$$(M \models_\phi p(x,y)) \wedge (\forall\phi_x \in \Phi_D)((M \models_{\phi_x} Q(x)) \wedge (\forall\phi_{xx} \in \Phi_D)(M \models_{\phi_{xx}} (q(x) \wedge R(x,z))))$$
$$\Longrightarrow$$
$$(M \models_\phi p(x,y)) \wedge (\forall\phi_x \in \Phi_D)((M \models_{\phi_x} Q(x)) \wedge (M \models_{\phi_x} \forall x(q(x) \wedge R(x,z))))$$
$$\Longrightarrow$$
$$(M \models_\phi p(x,y)) \wedge (\forall\phi_x \in \Phi_D)(M \models_{\phi_x} (Q(x) \wedge \forall x(q(x) \wedge R(x,z))))$$
$$\Longrightarrow$$
$$(M \models_\phi p(x,y)) \wedge M \models_\phi \forall x(Q(x) \wedge \forall x(q(x) \wedge R(x,z)))$$
$$\Longrightarrow$$
$$M \models_\phi (p(x,y) \wedge \forall x(Q(x) \wedge \forall x(q(x) \wedge R(x,z))))$$

We can show that a m-wff is true if and only if an atomic m-wff of its translation is true.

Proposition 4 *Let M be an interpretation and $\mathcal{A}_\phi$ be a m-wff with the taf ϕ and $M \models_\phi A$ be an atomic m-wff from $\mathcal{A}_\phi$ by the translation 1. Then $\mathcal{A}_\phi$ is true if and only $M \models_\phi A$ is true.*

Proof. We prove the above by induction on the number r of connectives and quantifiers in $\mathcal{A}_\phi$. Assume the result holds for all integers $< r$.

1. $\mathcal{A}_\phi$ is of the form $M \models_\phi A$ where A is a wff. This case is trivial.

2. $\mathcal{A}_\phi$ is of the form $\neg\mathcal{B}_\phi$. It is translated into $M \models_\phi \neg B$, where $\mathcal{B}_\phi$ is translated into $M \models_\phi B$. By the inductive hypothesis, $\mathcal{B}_\phi$ is true if and only if $M \models_\phi B$ is true. Therefore, $\neg\mathcal{B}_\phi$ is true if and only if $\neg M \models_\phi B$ is true if and only if $M \models_\phi \neg B$ is true.

3. $\mathcal{A}_\phi$ is of the form $\mathcal{B}_\phi \supset \mathcal{C}_\phi$. It is translated into $M \models_\phi B \supset C$, where $\mathcal{B}_\phi$ is translated into $M \models_\phi B$ and $\mathcal{C}_\phi$ is translated into $M \models_\phi C$. By the inductive hypothesis, $\mathcal{B}_\phi$ is true if and only if $M \models_\phi B$ is true, and $\mathcal{C}_\phi$ is true if and only if $M \models_\phi C$ is true. Therefore, $\mathcal{B} \supset \mathcal{C}$ is true if and only if $(M \models_\phi B) \supset (M \models_\phi C)$ is true if and only if $M \models_\phi B \supset C$.

4. $\mathcal{A}_\phi$ is of the form $(\forall \phi_v \in \Phi_D)\mathcal{B}_{\phi_v}$. It is translated into $M \models_\phi \forall v B$, where $\mathcal{B}_{\phi_v}$ is translated into $M \models_{\phi_v} B$. By the inductive hypothesis, $\mathcal{B}_{\phi_v}$ is true if and only if $M \models_{\phi_v} B$ is true . Therefore, $(\forall \phi_v \in \Phi_D)\mathcal{B}$ is true if and only if $(\forall \phi_v \in \Phi_D) M \models_{\phi_v} B$ is true if and only if $M \models_\phi \forall v B$ is true. □

Note that for every wff A, there exists a m-wff such that its translation is an atomic m-wff $M \models_\phi A$ because that atomic m-wff itself is a m-wff.

4 Relations over interpretations

Let M and M' be interpretations. M and M' are *comparable* with respect to a tuple of predicate constants $\mathbf{P}$ if and only if the following conditions are satisfied.

1. M and M' have the same domain D.

2. For every individual constant a, $(a)^M = (a)^{M'}$.

3. For every function constant f, $(f)^M = (f)^{M'}$.

4. For every predicate constant Q not in $\mathbf{P}$, $(Q)^M = (Q)^{M'}$.

Proposition 5 *Let interpretations M and M' with the domain D be comparable with respect to a tuple of predicate $\mathbf{P}$, and t be a term all of whose free variables are in $< x_1, ..., x_n >$, and ϕ and ϕ' be assignment functions which agree on the assignment of every x_i. Then, $\phi^M(t) = \phi'^{M'}(t)$.*

Proof. Since M and M' are comparable to $\mathbf{P}$, for every individual constant a, $(a)^M = (a)^{M'}$ and for every function constant f, $(f)^M = (f)^{M'}$. Every free variable x in t, $\phi(x) = \phi'(x)$. We prove the above by induction on the number m of function letter in t. Assume the result holds for all integers $< m$.

1. If t is an individual constant a, then $\phi^M(a) = (a)^M = (a)^{M'} = \phi'^{M'}(a)$.

2. If t is an individual variable x, then $\phi^M(x) = \phi(x) = \phi'^{M'}(x)$.

3. If t is of the form $f^n(t_1, ..., t_n)$. Each t_i has fewer than m function letters. By inductive hypothesis, $\phi^M(t_i) = \phi'^{M'}(t_i)$. Then $\phi^M(f^n(t_1, ..., t_n)) = (f^n)^M(\phi^M(t_1), ..., \phi^M(t_n)) = (f^n)^{M'}(\phi'^{M'}(t_1), ..., \phi'^{M'}(t_n)) = \phi'^{M'}(f^n(t_1, ..., t_n))$. □

A predicate variable and a predicate constant are *similar* if and only if they have the same arity. A tuple of predicate variables $\mathbf{p}(=< p_1, ..., p_n >)$ and a tuple of predicate constants $\mathbf{P}(=< P_1, ..., P_n >)$ are *similar* (or we say $\mathbf{p}$ is similar to $\mathbf{P}$) if and only if each variable p_i of $\mathbf{p}$ and each corresponding constant P_i of $\mathbf{P}$ are similar. We write a wff A with some of the predicate constants $P_1, ..., P_n$ in a tuple of predicate constants $\mathbf{P}$ as $A(\mathbf{P})$. Then we write as $A(\mathbf{p})$ the result of substituting in A the predicate variables $p_1, ..., p_n$ for all occurrences of $P_1, ..., P_n$ respectively.

Proposition 6 *Let* $\mathbf{p}$ *and* $\mathbf{P}$ *be similar and* $A(\mathbf{P})$ *be a formula which does not contain any predicate variable in* $\mathbf{p}$ *and* M *and* M' *be a comparable interpretations with the domain* D *and* ϕ *be an assignment function. If for every* P_i *in* $\mathbf{P}$ *and* p_i *in* $\mathbf{p}$, $(P_i)^{M'} = \phi(p_i)$, *then* $M' \models_\phi A(\mathbf{P})$ *if and only if* $M \models_\phi A(\mathbf{p})$.

Proof. We prove the above by induction on the number r of connectives and quantifiers in A. Assume the result holds for all integers $< r$.

1. $A(\mathbf{P})$ is of the form $P^n(t_1, ..., t_n)$ where P^n is a n-ary predicate constants not in $\mathbf{P}$. Then $(P^n)^{M'} = (P^n)^M$. By the proposition 5, for each t_i, $\phi^{M'}(t_i) = \phi^M(t_i)$. Therefore, $M' \models_\phi P^n(t_1, ..., t_n)$ if and only if $< \phi^{M'}(t_1), ..., \phi^{M'}(t_n) > \in (P^n)^{M'}$ if and only if $< \phi^M(t_1), ..., \phi^M(t_n) > \in (P^n)^M$ if and only if $M \models_\phi P^n(t_1, ..., t_n)$

2. $A(\mathbf{P})$ is of the form $P^n(t_1, ..., t_n)$ where P^n is a n-ary predicate constants in $\mathbf{P}$. Let p^n be a corresponding variable with P^n. Then $(P^n)^{M'} = \phi(p^n)$. By the proposition 5, for each t_i, $\phi^{M'}(t_i) = \phi^M(t_i)$. Therefore, $M' \models_\phi P^n(t_1, ..., t_n)$ if and only if $< \phi^{M'}(t_1), ..., \phi^{M'}(t_n) > \in (P^n)^{M'}$ if and only if $< \phi^M(t_1), ..., \phi^M(t_n) > \in \phi(p^n)$ if and only if $M \models_\phi p^n(t_1, ..., t_n)$

3. $A(\mathbf{P})$ is of the form $p^n(t_1, ..., t_n)$ where p^n is a n-ary predicate variables. p^n is not in $\mathbf{p}$ because of the condition of $A(\mathbf{P})$. Thus $\phi^{M'}(p^n) = \phi^M(p^n)$. By the proposition 5, for each t_i, $\phi^{M'}(t_i) = \phi^M(t_i)$. Therefore $M' \models_\phi p^n(t_1, ..., t_n)$ if and only if $< \phi^{M'}(t_1), ..., \phi^{M'}(t_n) > \in \phi(p^n)$ if and only if $< \phi^M(t_1), ..., \phi^M(t_n) > \in \phi(p^n)$ if and only if $M \models_\phi p^n(t_1, ..., t_n)$

4. $A(\mathbf{P})$ is of the form $\neg B(\mathbf{P})$. $B(\mathbf{P})$ has fewer than r connectives and quantifications and $B(\mathbf{P})$ does not contain any predicate variable in $\mathbf{p}$. By the inductive hypothesis, $M' \models_\phi B(\mathbf{P})$ if and only if $M \models_\phi B(\mathbf{p})$. By the definition of the satisfaction, $M' \models_\phi \neg B(\mathbf{P})$ if and only if $\neg M' \models_\phi B(\mathbf{P})$, and $M \models_\phi \neg B(\mathbf{p})$ if and only if $\neg M \models_\phi B(\mathbf{p})$. Therefore $M' \models_\phi \neg B(\mathbf{P})$ if and only if $M \models_\phi \neg B(\mathbf{p})$.

5. $A(\mathbf{P})$ is of the form $B(\mathbf{P}) \supset C(\mathbf{P})$. This case is proved in a similar way to the previous case.

6. $A(\mathbf{P})$ is of the form $\forall v B(\mathbf{P})$ where v is a variable. Assume $M' \models_\phi A$. Then for every ϕ_v in Φ_D, $M' \models_{\phi_v} B(\mathbf{P})$. And $B(\mathbf{P})$ does not contain any variable in $\mathbf{p}$. By the inductive hypothesis that $M' \models_{\phi_v} B(\mathbf{P})$ if and only if $M \models_{\phi_v} B(\mathbf{p})$. Hence $M \models_\phi \forall v B(\mathbf{p})$. The converse also holds in a similar way. □

Let M and M' be comparable interpretations with respect to $\mathbf{P}$. We define meta-relation over comparable interpretations by extending the notion of meta-formula.

Definition 5 *Meta-relation*

1. If A is a wff and ϕ is an assignment function, $M \models_\phi A$, $M' \models_\phi A$ is a meta-relation called an *atomic meta-relation.*

2. If $\mathcal{A}$ and $\mathcal{B}$ are meta-relations, ϕ is an assignment function and v is a variable, then $\neg\mathcal{A}$, $\mathcal{A} \supset \mathcal{B}$, $(\forall \phi_v \in \Phi_D)\mathcal{A}$ are meta-relations.

3. An expression is a meta-relation only if it satisfies one of the above conditions.

We can define a *subformula* of a meta-relation, a *quantified meta-relation*, a *scope* of a quantifier, a *depth* of subformula of a meta-relation and a *depth* of a scope in a similar way to the meta-relation. Then we define the particular meta-relations which can be translated to the wff. We call them *well-defined relations (wdrs).*

Definition 6 *Well-defined relations (wdrs)*

A meta-relation $\mathcal{A}$ is a wdr if and only if the following conditions are satisfied.

1. Every atomic meta-relation in $\mathcal{A}$ whose depth is 0 has the same assignment function ϕ. We call this ϕ the *toplevel assignment function(taf).* We write as $\mathcal{A}_\phi$ a wdr with a taf ϕ.

2. For every quantified meta-relation $(\forall \phi'_v \in \Phi_D)\mathcal{B}$ whose depth is 0, ϕ'_v is different from the taf ϕ at most in the assignment of a variable v, that is $\phi'_v = \phi_v$.

3. In every scope of the quantifier "$(\forall \phi_v \in \Phi_D)$" in $\mathcal{A}$ whose depth is m, every atomic meta-relation in the scope whose depth is m has the same assignment function ϕ_v in the quantifier. We call this ϕ_v the *assignment function of the scope.*

4. In every scope of the quantifier "$(\forall \phi_v \in \Phi_D)$" in $\mathcal{A}$ whose depth is m and whose assignment function is ϕ_v, for every quantified meta-relation in the scope $(\forall \phi'_u \in \Phi_D)\mathcal{B}$ whose depth is m, ϕ'_u is different from ϕ_v at most in the assignment of a variable u, that is, $\phi'_u = \phi_{vu}$.

Note that a subformula $\mathcal{B}$ of a wdr $\mathcal{A}$ is also a wdr. We sometimes write as $\mathcal{B}_{\phi'}$ a subformula of a wdr with a taf ϕ'. (For example, if $\mathcal{A}_\phi$ is of the form $\neg\mathcal{B}$, we also write $\neg\mathcal{B}_\phi$) From now on, we write a wdr over M and M' with the taf ϕ as $\mathcal{R}(M, M')_\phi$.

We will later use the following proposition.

Proposition 7 *Let M' and M be a comparable with respect to* $\mathbf{P}$ *and* $\mathcal{R}(M',M)_\phi$ *be a wdr. Let* $\mathbf{p}$ *be similar to* $\mathbf{P}$ *such that every variables in $p_1,...,p_n$ is not contained in* $\mathcal{R}(M',M)_\phi$*. Then* $\mathcal{R}(M',M)_\phi$ *is true if and only if* $\mathcal{R}(M',M)_{\phi_{p_1...p_n}}$ *is true.*

Proof. We prove the above by induction on the number r of connectives and quantifiers in $\mathcal{R}(M',M)$. Assume the result holds for all integers $< r$.

1. $\mathcal{R}(M',M)_\phi$ is of the form $M \models_\phi A$ Since A does not contain any variable in $\mathbf{p}$, ϕ and $\phi_{p_1...p_n}$ agree on the assignments of free variables in A. By the proposition 2, $M \models_\phi A$ is true if and only if $M \vdash_{\phi_{p_1...p_n}}$ is true.

2. $\mathcal{R}(M',M)_\phi$ is of the form $M' \models_\phi A$. This case is proved in a similar way to the previous case.

3. $\mathcal{R}(M',M)_\phi$ is of the form $\neg\mathcal{A}(M',M)_\phi$. By the inductive hypothesis, $\mathcal{A}(M',M)_\phi$ is true if and only if $\mathcal{A}(M',M)_{\phi_{p_1...p_n}}$ is true. Therefore $\neg\mathcal{A}(M',M)_\phi$ is true if and only if $\neg\mathcal{A}(M',M)_{\phi_{p_1...p_n}}$ is true.

4. $\mathcal{R}(M',M)_\phi$ is of the form $\mathcal{A}(M',M)_\phi \supset \mathcal{B}(M',M)_\phi$. This case is proved in a similar way to the previous case.

5. $\mathcal{R}(M',M)_\phi$ is of the form $(\forall\phi_v \in \Phi_D)\mathcal{A}(M',M)_{\phi_v}$. Assume $(\forall\phi_v \in \Phi_D)\mathcal{A}(M',M)_{\phi_v}$. Then, for every ϕ_v in Φ_D, $\mathcal{A}(M',M)_{\phi_v}$ is true. Take any $\phi_{p_1...p_nv}$. Then we can take ϕ_v such that $\phi_v(v) = \phi_{p_1...p_nv}(v)$. Since v is not in $v_1,...,v_n$, $\phi_{p_1...p_nv} = \phi_{vp_1...p_n}$ By the inductive hypothesis, $\mathcal{A}(M',M)_{\phi_v}$ is true if and only if $\mathcal{A}(M',M)_{\phi_{vp_1...p_n}}$ is true if and only if $\mathcal{A}(M',M)_{\phi_{p_1...p_nv}}$ is true. Since for every ϕ_v in Φ_D, $\mathcal{A}(M',M)_{\phi_v}$ is true, for every $\phi_{p_1...p_nv}$ in Φ_D, $\mathcal{A}(M',M)_{\phi_{p_1...p_nv}}$ is true. Hence, $(\forall\phi_{p_1...p_nv}\in\Phi_D)\mathcal{A}(M',M)_{\phi_{p_1...p_nv}}$ is true. □

Example 3 *Well-defined relations (wdrs)*

- $(\forall\phi_x\in\Phi_D)(M \models_\phi P(x) \wedge M' \models_{\phi_x} Q(x))$ is a meta-relation but not a wdr because this formula violates condition 2.
- $M' \models_\phi P(x) \wedge (\forall\phi_{xy} \in \Phi_D)(M \models_{\phi_{xy}} Q(x))$ is a meta-relation but not a wdr because this formula violates condition 3.
- $M' \models_\phi p(x,y) \wedge (\forall\phi_x\in\Phi_D)(M \models_{\phi_x} Q(x) \wedge (\forall\phi_{xz}\in\Phi_D)(M' \models_{\phi_{xz}} (q(x) \wedge R(x,z))))$ is a wdr.

By the definition of wdr, we can convert any wdr into a m-wff of M by the following translation.

Translation 2: from a wdr to a m-wff
Let M' and M be a comparable with respect to $\mathbf{P}$ and $\mathcal{R}(M',M)_\phi$ be a wdr. Let $\mathbf{p}$ be similar to $\mathbf{P}$ such that every predicate variable in $\mathbf{p}$ is not contained in $\mathcal{R}(M',M)_\phi$.

1. $\mathcal{R}(M',M)_\phi$ is the form of $M \models_\phi A$. It is translated into itself.

2. $\mathcal{R}(M', M)_\phi$ is the form of $M' \models_\phi A(\mathbf{P})$. It is translated into $M \models_\phi A(\mathbf{p})$.

3. $\mathcal{R}(M', M)_\phi$ is the form of $\neg\mathcal{A}(M', M)_\phi$. It is translated into $\neg\mathcal{A}_\phi$, where $\mathcal{A}(M', M)_\phi$ is translated into $\mathcal{A}_\phi$.

4. $\mathcal{R}(M', M)_\phi$ is the form of $\mathcal{A}(M', M)_\phi \supset \mathcal{B}(M', M)_\phi$. It is translated into $\mathcal{A}_\phi \supset \mathcal{B}_\phi$, where $\mathcal{A}(M', M)_\phi$ is translated into $\mathcal{A}_\phi$, and $\mathcal{B}(M', M)_\phi$ is translated into $\mathcal{B}_\phi$.

5. $\mathcal{R}(M', M)_\phi$ is of the form $(\forall\phi_v \in \Phi_D)\mathcal{A}(M', M)_{\phi_v}$. It is translated into $(\forall\phi_v \in \Phi_D)\mathcal{A}_{\phi_v}$, where $\mathcal{A}(M', M)_{\phi_v}$ is translated into $\mathcal{A}_{\phi_v}$.

Example 4 *Translation from a wdr into a m-wff*

Let M and M' be comparable with respect to $< P, Q >$ and $< p, q >$ be similar to $< P, Q >$ and ϕ be an assignment function.
Let $\mathcal{R}(M, M')_\phi$ be:

$$(M' \models_\phi P(x, y)) \wedge (\forall\phi_x \in \Phi_D)((M \models_{\phi_x} Q(x)) \wedge (\forall\phi_{xx} \in \Phi_D)(M' \models_{\phi_{xx}} (Q(x) \wedge R(x, z))))$$

$$\Longrightarrow$$

$$(M \models_\phi p(x, y)) \wedge (\forall\phi_x \in \Phi_D)((M \models_{\phi_x} Q(x)) \wedge (\forall\phi_{xx} \in \Phi_D)(M \models_{\phi_{xx}} (q(x) \wedge R(x, z))))$$

If ϕ satisfies the following condition in the proposition 8, we can show that a wdr is true if and only if a m-wff of its translation is true.

Proposition 8 *Let M' and M be a comparable with respect to* $\mathbf{P}$ *and $\mathcal{R}(M', M)_\phi$ be a wdr. Let* $\mathbf{p}$ *be similar to* $\mathbf{P}$ *such that every predicate variable in* $\mathbf{p}$ *is not contained in $\mathcal{R}(M', M)_\phi$. Let $\mathcal{R}_\phi$ be a m-wff from $\mathcal{R}(M', M)_\phi$ by the translation 2. If for every P_i in* $\mathbf{P}$ *and p_i in* $\mathbf{p}$, *$\phi(p_i) = (P_i)^{M'}$, then $\mathcal{R}(M', M)_\phi$ is true if and only if $\mathcal{R}_\phi$ is true.*

Proof. We prove the above by induction on the number r of connectives and quantifiers in $\mathcal{R}(M', M)_\phi$. Assume the result holds for all integers $< r$.

1. $\mathcal{R}(M', M)_\phi$ is of the form $M \models_\phi A$. This case is trivial.

2. $\mathcal{R}(M', M)_\phi$ is of the form $M' \models_\phi A(\mathbf{P})$. It is translated into $M \models_\phi A(\mathbf{p})$. $A(\mathbf{P})$ does not contain any predicate variable in $\mathbf{p}$ and for every p_i in $\mathbf{p}$ and corresponding P_i in $\mathbf{P}$, $\phi(p_i) = (P_i)^{M'}$. By the proposition 6, $M' \models_\phi A(\mathbf{P})$ is true if and only if $M \models_\phi A(\mathbf{p})$ is true.

3. $\mathcal{R}(M', M)_\phi$ is of the form $\neg\mathcal{A}(M', M)_\phi$. It is translated into $\neg\mathcal{A}_\phi$, where $\mathcal{A}(M', M)_\phi$ is translated into $\mathcal{A}_\phi$. By the inductive hypothesis, $\mathcal{A}(M', M)_\phi$ is true if and only if $\mathcal{A}_\phi$ is true. Therefore $\neg\mathcal{A}(M', M)_\phi$ is true if and only if $\neg\mathcal{A}_\phi$ is true.

4. $\mathcal{R}(M', M)_\phi$ is of the form $\mathcal{A}(M', M)_\phi \supset \mathcal{B}(M', M)_\phi$. It is translated into $\mathcal{A}_\phi \supset \mathcal{B}_\phi$, where $\mathcal{A}(M', M)_\phi$ is translated into $\mathcal{A}_\phi$, and $\mathcal{B}(M', M)_\phi$ is translated into $\mathcal{B}_\phi$. By the inductive hypothesis, $\mathcal{A}(M', M)_\phi$ is true if and only if $\mathcal{A}_\phi$ is true, and $\mathcal{B}(M', M)_\phi$ is true if and only if $\mathcal{B}_\phi$ is true. Therefore, $\mathcal{A}(M', M)_\phi \supset \mathcal{B}(M', M)_\phi$ is true if and only if $\mathcal{A}_\phi \supset \mathcal{B}_\phi$ is true.

5. $\mathcal{R}(M',M)_\phi$ is of the form $(\forall\phi_v\in\Phi_D)\mathcal{A}(M',M)_{\phi_v}$. It is translated into $(\forall\phi_v\in\Phi_D)\mathcal{A}_{\phi_v}$, where $\mathcal{A}(M',M)_{\phi_v}$ is translated into $\mathcal{A}_{\phi_v}$. By the inductive hypothesis, $\mathcal{A}(M',M)_{\phi_v}$ is true if and only if $\mathcal{A}_{\phi_v}$ is true. Therefore, $(\forall\phi_v\in\Phi_D)\mathcal{A}(M',M)_{\phi_v}$ is true if and only if $(\forall\phi_v\in\Phi_D)\mathcal{A}_{\phi_v}$ is true. □

Note that we can covert any wdr into an atomic m-wff by the translation 1 and 2.

Example 5 *Translation from a wdr into an atomic w-wff*
Let M and M' are comparable with respect to $< P, Q >$ and $< p, q >$ be similar to $< P, Q >$ and ϕ be an assignment function.
Let $\mathcal{R}(M,M')_\phi$ be:
$(M' \models_\phi P(x,y)) \wedge (\forall\phi_x\in\Phi_D)((M \models_{\phi_x} Q(x)) \wedge (\forall\phi_{xz}\in\Phi_D)(M' \models_{\phi_{xz}} (Q(x) \wedge R(x,z))))$
$\Longrightarrow$ (see example 4)
$(M \models_\phi p(x,y)) \wedge (\forall\phi_x\in\Phi_D)((M \models_{\phi_x} Q(x)) \wedge (\forall\phi_{xz}\in\Phi_D)(M \models_{\phi_{xz}} (q(x) \wedge R(x,z))))$
$\Longrightarrow$ (see example 2)
$(M \models_\phi p(x,y)) \wedge \forall x(Q(x) \wedge \forall x(q(x) \wedge R(x,z)))$

Actually, for any wff $A(\mathbf{p})$, there exists a wdr such that its translation is an atomic m-wff $M \models_\phi A(\mathbf{p})$ because a wdr $M' \models_\phi A(\mathbf{P})$ is translated into that atomic m-wff.

We can show that if an assignment function ϕ satisfies the condition in the proposition 8, a wdr is true if and only if an atomic m-wff of its translation is true.

Lemma 1 *Let M' and M be a comparable with respect to $\mathbf{P}$ and $\mathcal{R}(M',M)_\phi$ be a wdr. Let $\mathbf{p}$ be similar to $\mathbf{P}$ such that every predicate variable in $\mathbf{p}$ is not contained in $\mathcal{R}(M',M)_\phi$. Let $\mathcal{R}_\phi$ be a m-wff from $\mathcal{R}(M',M)_\phi$ by the translation 2. And let $M \models_\phi R(\mathbf{p})$ be an atomic m-wff from $\mathcal{R}_\phi$ by the translation 1. If for every P_i in $\mathbf{P}$ and p_i in $\mathbf{p}$, $\phi(p_i) = (P_i)^{M'}$, then $\mathcal{R}(M',M)_\phi$ is true if and only if $M \models_\phi R(\mathbf{p})$ is true.*

Proof. Since for every P_i in $\mathbf{P}$ and p_i in $\mathbf{p}$, $\phi(p_i) = (P_i)^{M'}$. By the proposition 8, $\mathcal{R}(M',M)_\phi$ is true if and only if $\mathcal{R}_\phi$ is true. And by the proposition 4, $\mathcal{R}_\phi$ is true if and only if $M \models_\phi R(\mathbf{p})$ is true. Therefore, $\mathcal{R}(M',M)_\phi$ is true if and only if $M \models_\phi R(\mathbf{p})$ is true. □

Example 6 *Equivalence between a wdr and an atomic m-wff*
Let M and M' are comparable with respect to $< P, Q >$ and $< p, q >$ be similar to $< P, Q >$ and ϕ be an assignment function such that $\phi(p) = (P)^{M'}$ and $\phi(q) = (Q)^{M'}$.
Let $\mathcal{R}(M,M')_\phi$ be:
$(M' \models_\phi P(x,y)) \wedge (\forall\phi_x\in\Phi_D)((M \models_{\phi_x} Q(x)) \wedge (\forall\phi_{xz}\in\Phi_D)(M' \models_{\phi_{xz}} (Q(x) \wedge R(x,z))))$.
It is true if and only if
$(M \models_\phi p(x,y)) \wedge (\forall\phi_x\in\Phi_D)((M \models_{\phi_x} Q(x)) \wedge (\forall\phi_{xz}\in\Phi_D)(M \models_{\phi_{xz}} (q(x) \wedge R(x,z))))$
is true
if and only if
$M \models_\phi p(x,y) \wedge \forall x(Q(x) \wedge \forall x(q(x) \wedge R(x,z)))$ is true.

Now we prove the following theorem closely related to a link between minimal models in preference order and 2nd-order wff.

Theorem 1 *Let M' and M be models with the domain D which are comparable with respect to $\mathbf{P}$ and $\mathcal{R}(M', M)_\phi$ be a wdr. Let $\mathbf{p}$ be similar to $\mathbf{P}$ such that every predicate variable in $\mathbf{p}$ is not contained in $\mathcal{R}(M', M)_\phi$. Let its translation using $\mathbf{p}$ be $M \models_\phi R(\mathbf{p})$. There exists M' such that $R(M', M)_\phi$ is true if and only if $M \models_\phi \exists p_1 ... \exists p_n R(\mathbf{p})$ is true.*

Proof. Assume there exists M' such that $R(M', M)_\phi$ is true. Since $R(M', M)_\phi$ does not contain any variable in $\mathbf{p}$, for every $\phi_{p_1 ... p_n}$ in Φ_D, $R(M', M)_{\phi_{p_1 ... p_n}}$ is true by the proposition 7. Let for every p_i in $\mathbf{p}$ and P_i in $\mathbf{P}$, $\phi_{p_1 ... p_n}(p_i) = (P_i)^{M'}$. Then, by the lemma 1, $R(M', M)_{\phi_{p_1 ... p_n}}$ is true if and only if $M \models_{\phi_{p_1 ... p_n}} R(\mathbf{p})$ is true. Then, by the proposition 3, $M \models_\phi \exists p_1 ... \exists p_n R(\mathbf{p})$ is true.
Assume $M \models_\phi \exists p_1 ... \exists p_n R(\mathbf{p})$ is true. By the proposition 3, $M \models_{\phi_{p_1 ... p_n}} R(\mathbf{p})$ is true. There exists M' such that M' and M are comparable with respect to $\mathbf{P}$, and for every P_i in $\mathbf{P}$ and p_i in $\mathbf{p}$, $(P_i)^{M'} = \phi_{p_1 ... p_n}(p_i)$. Then, by the lemma 1, $M \models_{\phi_{p_1 ... p_n}} R(\mathbf{p})$ is true if and only if $R(M', M)_{\phi_{p_1 ... p_n}}$ is true. Then, by the proposition 7, $R(M', M)_\phi$ is true. □

Corollary 1 *Let M' and M be models with the domain D which are comparable with respect to $\mathbf{P}$ and $\mathcal{R}(M', M)_\phi$ be a wdr. Let $\mathbf{p}$ be similar to $\mathbf{P}$ such that every predicate variable in $\mathbf{p}$ is not contained in $\mathcal{R}(M', M)_\phi$. Let its translation using $\mathbf{p}$ be $M \models_\phi R(\mathbf{p})$. Then the followings are equivalent.*

1. *For every assignment function ϕ in Φ_D and for every model M', $\neg\mathcal{R}(M', M)_\phi$ is true.*

2. *$M \models \neg\exists p_1 ... \exists p_n R(\mathbf{p})$ is true.*

Proof. For every M', $\neg\mathcal{R}(M', M)_\phi$ is true if and only if not $(M \models_\phi \exists p_1 ... \exists p_n R(\mathbf{p}))$ is true. by the theorem 1. It is true if and only if $M \models_\phi \neg\exists p_1 ... \exists p_n R(\mathbf{p})$ is true. Therefore, for every ϕ, the followings are equivalent.

1. For every M', $\neg\mathcal{R}(M', M)_\phi$ is true.

2. $M \models_\phi \neg\exists p_1 ... \exists p_n R(\mathbf{p})$ is true.

Therefore, the followings are equivalent.

1. For every ϕ and for every M', $\neg\mathcal{R}(M', M)_\phi$ is true.

2. For every ϕ, $M \models_\phi \neg\exists p_1 ... \exists p_n R(\mathbf{p})$ is true.

The condition 2 is true if and only if $M \models \neg\exists p_1 ... \exists p_n R(\mathbf{p})$ is true. □

Corollary 2 *Let M' and M be models with the domain D which are comparable with respect to $\mathbf{P}$ and $\mathcal{R}(M', M)_\phi$ be a wdr and $A(\mathbf{P})$ be a wff. Let $\mathbf{p}$ be similar to $\mathbf{P}$ such that every predicate variable in $\mathbf{p}$ is not contained in $\mathcal{R}(M', M)_\phi$ and $A(\mathbf{P})$. Let the translation of $\mathcal{R}(M', M)_\phi$ using $\mathbf{p}$ be $M \models_\phi R(\mathbf{p})$. The followings are equivalent.*

1. M is a model of $A(\mathbf{P})$ and for every assignment function ϕ in Φ_D and for every M', $((M' \models_\phi A(\mathbf{P})) \supset \neg\mathcal{R}(M', M)_\phi)$ is true.

2. $M \models A(\mathbf{P}) \wedge \neg\exists p_1...\exists p_n(A(\mathrm{p}) \wedge R(\mathrm{p}))$ is true.

Proof. Assume condition 1. M is a model of $A(\mathbf{P})$ if and only if $M \models A(\mathbf{P})$. Since $(M' \models_\phi A(\mathbf{P})) \wedge \mathcal{R}(M', M)_\phi$ is a wdr, we can translate it into an atomic m-wff. By the translation 2, the above wdr becomes:
$(M \models_\phi A(\mathrm{p})) \wedge \mathcal{R}_\phi$,
where $\mathcal{R}_\phi$ is a translated m-wff from $R(M', M)_\phi$.
Then by the translation 1, the above m-wff becomes:
$M \models_\phi A(\mathrm{p}) \wedge R(\mathrm{p})$.
Therefore, by the corollary 1, for every assignment function ϕ in Φ_D and for every model M', $\neg((M' \models_\phi A(\mathbf{P})) \wedge \mathcal{R}(M', M)_\phi)$ is true if and only if $M \models \neg\exists p_1...\exists p_n(A(\mathrm{p}) \wedge R(\mathrm{p}))$ is true. Therefore, condition 1 is true if and only if $M \models A(\mathbf{P})$ and $M \models \neg\exists p_1...\exists p_n(A(\mathrm{p}) \wedge R(\mathrm{p}))$ are true. They are true if and only if condition 2 is true. □

If $A(\mathbf{P})$ is a closed wff and some assignment function ϕ in Φ_D, $M' \models_\phi A(\mathbf{P})$ is true, then M' is a model of $A(\mathbf{P})$ by the proposition 2. Therefore if $A(\mathbf{P})$ is closed then condition 1 of the corollary 2 becomes the following.

1' *M is a model of $A(\mathbf{P})$ and for every M', if M' is a model of $A(\mathbf{P})$ then for every assignment function ϕ, $\neg\mathcal{R}(M', M)_\phi$ is true.*

Note that $A(\mathbf{P})$ must be closed if the condition 1 of the corollary 2 and the condition 1' is equivalent. If $A(\mathbf{P})$ is not closed, then the condition 1 implies the condition 1' but the converse is not true. We can show a counter-example. Let M' and M be models with the domain D which are comparable with respect to P and $A(\mathbf{P})$ be $P(x)$ and $\mathcal{R}(M', M)_\phi$ be $M' \models_\phi \neg P(a)$ and $(P)^M = D$. Then the condition 1' becomes the following condition that for every ϕ in Φ_D, $M \models_\phi P(x)$ and for every M', if for every ϕ in Φ_D, $M' \models_\phi P(x)$ then for every ϕ in Φ_D, $\neg M' \models_\phi \neg P(a)$ $(= (M' \models_\phi P(a)))$ is true. We can easily see that this condition is true. Let $\phi(p) = \{(b)^M\}$ and $\phi(x) = (b)^M$ and $(a)^M \neq (b)^M$. Then, $M \models_\phi p(x) \wedge \neg p(a)$. Therefore, there exists ϕ such that $M \models_\phi p(x) \wedge \neg p(a)$. Therefore, condition 1, that is, $M \models P(x) \wedge \neg\exists p(p(x) \wedge \neg p(a))$ is false.

We say that $\mathcal{R}(M', M)$ is a *partial order relation* if and only if for every assignment function ϕ in Φ_D, the following conditions are true:

1. for every M, $\mathcal{R}(M, M)_\phi$ is false.

2. for every M, M' and M'', $\mathcal{R}_\phi(M'', M')$ and $\mathcal{R}_\phi(M', M)$ implies $\mathcal{R}_\phi(M'', M)$.

Let $A(\mathbf{P})$ be a closed wff and $\mathcal{R}(M', M)$ be a partial order relation. We say that M is a *minimal model* with respect to a closed wff $A(\mathbf{P})$ and a wdr $\mathcal{R}(M', M)$ if and only if condition 1' is satisfied. Then we obtain the following formal semantics of general circumscription.

Corollary 3 *M is a minimal model with respect to a closed wff $A(\mathbf{P})$ and a wdr $\mathcal{R}(M', M)$ if and only if M is a model of $A(\mathbf{P}) \wedge \neg\exists p_1 ... \exists p_n(A(\mathbf{p}) \wedge R(\mathbf{p}))$.*

Proof. By the above discussion and the corollary 2. □

In the next section, we show some applications of this framework to nonmonotonic reasoning.

5 Application to formalizing nonmonotonic reasoning

5.1 Formula Circumscription

Consider the following wdr.

Let M and M' be interpretations with the domain D which are comparable with respect to $\mathbf{P} =< P_1, ..., P_n >$. And let $E(\mathbf{P}, x_1, ..., x_n)$ be a wff in which P_i in $\mathbf{P}$ and individual constants $x_1, ..., x_n$ occur free.
$\mathcal{R}(M', M)_\phi$ if and only if

$$\begin{aligned}&\forall\phi_{x_1}\in\Phi_D...\forall\phi_{x_1x_2...x_n}\in\Phi_D((M' \models_{\phi_{x_1...x_n}} E(\mathbf{P}, x_1, ..., x_n)) \supset \\ &\qquad\qquad (M \models_{\phi_{x_1...x_n}} E(\mathbf{P}, x_1, ..., x_n)))\wedge \\ &\neg\forall\phi_{x_1}\in\Phi_D...\forall\phi_{x_1x_2...x_n}\in\Phi_D((M \models_{\phi_{x_1...x_n}} E(\mathbf{P}, x_1, ..., x_n)) \supset \\ &\qquad\qquad (M' \models_{\phi_{x_1...x_n}} E(\mathbf{P}, x_1, ..., x_n))).\end{aligned}$$

Note that this wdr is a partial order. Let $\mathbf{p} =< p_1, ..., p_n >$ be similar to $\mathbf{P}$. The corresponding atomic m-wff $M \models_\phi R(\mathbf{p})$ is as follows (by using translation 1 and 2).

$$\begin{aligned}M \models_\phi &\forall x_1...\forall x_n(E(\mathbf{p}, x_1, ..., x_n) \supset E(\mathbf{P}, x_1, ..., x_n))\wedge \\ &\neg\forall x_1...\forall x_n(E(\mathbf{P}, x_1, ..., x_n) \supset E(\mathbf{p}, x_1, ..., x_n)).\end{aligned}$$

A model of a closed wff $A(\mathbf{P})$, M, is a minimal model with respect to $\mathcal{R}(M', M)$ if and only if M satisfies the following wff.

$$M \models A(\mathbf{P}) \wedge \neg\exists p_1...\exists p_n(A(\mathbf{p}) \wedge R(\mathbf{p})).$$

The wff of the above atomic m-wff is the form of formula circumscription.

5.2 Parallel Circumscription

Consider the following wdr.

Let M and M' be interpretations with the domain D which are comparable with respect to $\mathbf{P} =< P_1, ..., P_n >$. And let P_i in $\mathbf{P}$ be a_i-ary predicate constant.
$\mathcal{R}(M', M)_\phi$ if and only if

$$\begin{aligned}&\forall\phi_{x_1}\in\Phi_D...\forall\phi_{x_1x_2...x_{a_1}}\in\Phi_D((M' \models_{\phi_{x_1...x_{a_1}}} P_1(x_1, ..., x_{a_1})) \supset \\ &\qquad\qquad (M \models_{\phi_{x_1...x_{a_1}}} P_1(x_1, ..., x_{a_1})))\wedge \\ &\qquad\qquad \vdots \\ &\forall\phi_{x_1}\in\Phi_D...\forall\phi_{x_1x_2...x_{a_n}}\in\Phi_D((M' \models_{\phi_{x_1...x_{a_n}}} P_n(x_1, ..., x_{a_n})) \supset \\ &\qquad\qquad (M \models_{\phi_{x_1...x_{a_n}}} P_n(x_1, ..., x_{a_n})))\wedge \\ &\neg(\forall\phi_{x_1}\in\Phi_D...\forall\phi_{x_1...x_{a_1}}\in\Phi_D((M \models_{\phi_{x_1...x_{a_1}}} P_1(x_1, ..., x_{a_1})) \supset\end{aligned}$$

$$(M' \models_{\phi_{x_1 \ldots x_{a_1}}} P_1(x_1, ..., x_{a_1})))\wedge$$
$$\vdots$$
$$\forall \phi_{x_1} \in \Phi_D \ldots \forall \phi_{x_1 \ldots x_{a_n}} \in \Phi_D((M \models_{\phi_{x_1 \ldots x_{a_n}}} P_n(x_1, ..., x_{a_n})) \supset (M' \models_{\phi_{x_1 \ldots x_{a_n}}} P_n(x_1, ..., x_{a_n})))).$$

Note that this wdr is a partial order. Let $\mathbf{p} =< p_1, ..., p_n >$ be similar to $\mathbf{P}$. The corresponding atomic m-wff $M \models_\phi R(\mathbf{p})$ is as follows (by using translation 1 and 2).

$$\begin{aligned} M \models_\phi & \forall x_1 \ldots \forall x_{a_1}(p_1(x_1, ..., x_{a_1}) \supset P_1(x_1, ..., x_{a_1}))\wedge \\ & \vdots \\ & \forall x_1 \ldots \forall x_{a_n}(p_n(x_1, ..., x_{a_n}) \supset P_n(x_1, ..., x_{a_n}))\wedge \\ & \neg(\forall x_1 \ldots \forall x_{a_1}(P_1(x_1, ..., x_{a_1}) \supset p_1(x_1, ..., x_{a_1}))\wedge \\ & \vdots \\ & \forall x_1 \ldots \forall x_{a_n}(P_n(x_1, ..., x_{a_n}) \supset p_n(x_1, ..., x_{a_n}))). \end{aligned}$$

A model of a closed wff $A(\mathbf{P})$, M, is a minimal model with respect to $\mathcal{R}(M', M)$ if and only if M satisfies the following wff.

$$M \models A(\mathbf{P}) \wedge \neg \exists p_1 \ldots \exists p_n(A(\mathbf{p}) \wedge R(\mathbf{p})).$$

The wff of the above atomic m-wff is the form of parallel circumscription without variable predicates. If M and M' are comparable with respect to $\mathbf{P}' =< P_1, ..., P_n, P_{n+1}, ..., P_{n+m} >$, then considering the same form of $\mathcal{R}(M', M)$ gives circumscription with variable predicates $P_{n+1}, ..., P_{n+m}$.

5.3 Simple Default Reasoning

If we would like to express "most birds can fly", we do not need to introduce abnormal predicates minimized in circumscription but simply express that an interpretation M' is preferred to an interpretation M if and only if M' makes more birds fly than M.

Consider the following wdr. Let Fly and $Bird$ be unary predicate constants and M and M' be interpretations with the domain D which are comparable with respect to $< Fly >$.
$\mathcal{R}(M', M)_\phi$ if and only if
$$\forall \phi_x \in \Phi_D((M \models_{\phi_x} (Bird(x) \wedge Fly(x))) \supset (M' \models_{\phi_x} (Bird(x) \wedge Fly(x))))\wedge$$
$$\neg \forall \phi_x \in \Phi_D((M' \models_{\phi_x} (Bird(x) \wedge Fly(x))) \supset (M \models_{\phi_x} (Bird(x) \wedge Fly(x)))).$$

Note that this wdr is a partial order. Let fly be similar to Fly. The corresponding atomic m-wff $M \models_\phi R(fly)$ is as follows

$$\begin{aligned} M \models_\phi & \forall x((Bird(x) \wedge Fly(x)) \supset (Bird(x) \wedge fly(x)))\wedge \\ & \neg \forall x((Bird(x) \wedge fly(x)) \supset (Bird(x) \wedge Fly(x))). \end{aligned}$$

If we derive any result from $A(Fly) \wedge \neg \exists fly(A(fly) \wedge R(fly))$, it is true in every minimal models with respect to the above wdr. We show an example of derivation. Suppose $A(Fly) = Bird(t) \wedge \neg Fly(t)$ where t is an individual constants. Then the above wff becomes:

$$Bird(t) \wedge \neg Fly(t) \wedge \neg\exists fly(Bird(t) \wedge \neg fly(t) \wedge$$
$$\forall x((Bird(x) \wedge Fly(x)) \supset (Bird(x) \wedge fly(x))) \wedge$$
$$\neg\forall x((Bird(x) \wedge fly(x)) \supset (Bird(x) \wedge Fly(x))))$$

Suppose $fly = \lambda x(Fly(x) \vee (x \neq t))$. Then we can derive the following from the above wff.

$$Bird(t) \wedge \neg Fly(t) \wedge$$
$$\neg(Bird(t) \wedge \neg(Fly(t) \vee (t \neq t)) \wedge$$
$$\forall x((Bird(x) \wedge Fly(x)) \supset (Bird(x) \wedge (Fly(x) \vee (x \neq t)))) \wedge$$
$$\neg\forall x((Bird(x) \wedge (Fly(x) \vee (x \neq t))) \supset (Bird(x) \wedge Fly(x)))).$$

It is reduced to:

$$Bird(t) \wedge \neg Fly(t) \wedge \neg(\neg\forall x((x \neq t) \supset (Bird(x) \supset Fly(x)))),$$

which is equivalent to:

$$\forall x(x \neq t \equiv (Bird(x) \supset Fly(x))).$$

5.4 Prioritized Default Reasoning

If there are competing default rules such as "most birds can fly" and "most penguins can not fly", and we would like to give higher priority to the penguin rule, then we express these rules by an order that an interpretation M' is preferred to an interpretation M if and only if M' have more non-flying penguins than M and M' have more flying birds than M in the case that M' have same non-flying penguins as M.

Consider the following wdr. Let Fly, $Bird$ and $Penguin$ be unary predicate constants and M and M' be interpretations with the domain D which are compatible with respect to $< Fly >$.
$\mathcal{R}(M', M)_\phi$ if and only if

$$\forall\phi_x\in\Phi_D((M \models_{\phi_x} (Penguin(x) \wedge \neg Fly(x))) \supset$$
$$(M' \models_{\phi_x} (Penguin(x) \wedge \neg Fly(x)))) \wedge$$
$$(\forall\phi_x\in\Phi_D((M \models_{\phi_x} (Penguin(x) \wedge \neg Fly(x))) \equiv$$
$$(M' \models_{\phi_x} (Penguin(x) \wedge \neg Fly(x)))) \supset$$
$$\forall\phi_x\in\Phi_D((M \models_{\phi_x} (Bird(x) \wedge Fly(x))) \supset (M' \models_{\phi_x} (Bird(x) \wedge Fly(x))))) \wedge$$
$$\neg(\forall\phi_x\in\Phi_D((M' \models_{\phi_x} (Penguin(x) \wedge \neg Fly(x))) \supset$$
$$(M \models_{\phi_x} (Penguin(x) \wedge \neg Fly(x)))) \wedge$$
$$(\forall\phi_x\in\Phi_D((M' \models_{\phi_x} (Penguin(x) \wedge \neg Fly(x))) \equiv$$
$$(M \models_{\phi_x} (Penguin(x) \wedge \neg Fly(x)))) \supset$$
$$\forall\phi_x\in\Phi_D((M' \models_{\phi_x} (Bird(x) \wedge Fly(x))) \supset (M \models_{\phi_x} (Bird(x) \wedge Fly(x)))))).$$

Note that this wdr is a partial order. Let fly be similar to Fly. The corresponding atomic m-wff $M \models_\phi R(fly)$ is as follows

$$M \models_\phi \forall x((Penguin(x) \wedge \neg Fly(x)) \supset (Penguin(x) \wedge \neg fly(x))) \wedge$$
$$(\forall x((Penguin(x) \wedge \neg Fly(x)) \equiv (Penguin(x) \wedge \neg fly(x))) \supset$$
$$\forall x((Bird(x) \wedge Fly(x)) \supset (Bird(x) \wedge fly(x)))) \wedge$$

$$\neg(\forall x((Penguin(x) \wedge \neg fly(x)) \supset (Penguin(x) \wedge \neg Fly(x))) \wedge$$
$$(\forall x((Penguin(x) \wedge \neg fly(x)) \equiv (Penguin(x) \wedge \neg Fly(x))) \supset$$
$$\forall x((Bird(x) \wedge fly(x)) \supset (Bird(x) \wedge Fly(x))))).$$

If we derive any result from $A(Fly) \wedge \neg\exists fly(A(fly) \wedge R(fly))$, it is true in every minimal models with respect to the above wdr. We show an example of derivation. Suppose $A(Fly) = \forall x(Penguin(x) \supset Bird(x))$. Then the above wff becomes:

$$\forall x(Penguin(x) \supset Bird(x)) \wedge \neg\exists fly(\forall x(Penguin(x) \supset Bird(x)) \wedge R(fly)).$$

Suppose $fly = \lambda x(\neg Penguin(x))$. Then we can derive the following form the above wff.

$$\forall x(Penguin(x) \supset Bird(x)) \wedge$$
$$\neg(\forall x(Penguin(x) \supset Bird(x)) \wedge$$
$$\forall x((Penguin(x) \wedge \neg Fly(x)) \supset (Penguin(x) \wedge \neg(\neg Penguin(x)))) \wedge$$
$$(\forall x((Penguin(x) \wedge \neg Fly(x)) \equiv (Penguin(x) \wedge \neg(\neg Penguin(x)))) \supset$$
$$\forall x((Bird(x) \wedge Fly(x)) \supset (Bird(x) \wedge (\neg Penguin(x))))) \wedge$$
$$\neg(\forall x((Penguin(x) \wedge \neg(\neg Penguin(x))) \supset (Penguin(x) \wedge \neg Fly(x))) \wedge$$
$$(\forall x((Penguin(x) \wedge \neg(\neg Penguin(x))) \equiv (Penguin(x) \wedge \neg Fly(x))) \supset$$
$$\forall x((Bird(x) \wedge (\neg Penguin(x))) \supset (Bird(x) \wedge Fly(x)))))).$$

It is reduced to:

$$\forall x(Penguin(x) \supset Bird(x)) \wedge$$
$$\neg(\neg(\forall x(Penguin(x) \supset \neg Fly(x)) \wedge$$
$$(\forall x(Penguin(x) \supset \neg Fly(x)) \supset \forall x((Bird(x) \wedge \neg Penguin(x)) \supset Fly(x))))),$$

which is equivalent to:

$$\forall x(Penguin(x) \supset Bird(x)) \wedge \forall x(Penguin(x) \supset \neg Fly(x)) \wedge$$
$$\forall x((Bird(x) \wedge \neg Penguin(x)) \supset Fly(x)).$$

5.5 Minimal Change Models

We have presented a solution to the Yale shooting problem and multiple extension problem in the tree structured inheritance system in [Satoh87]. In that formalism, both types of reasoning are translated into reasoning in tree-structured multiple worlds and are regarded as selecting a preferred model which changes minimally in one direction. In the Yale shooting problem, the direction is from earlier state to later state. And in the inheritance systems, the direction is from superclass to subclass.

We briefly explain the formalism and show the ordering in a wdr. In the tree-structured world, we have two sorted-variables, *variables of properties* $p, p_1, p_2, \ldots$ and *variable of worlds* $w, w_1, w_2, \ldots$. And we introduce an individual constant 0 and a function *last* and two binary predicates T and $<$. 0 expresses the root of the tree and $last(w)$ gives a parent node of a world w and $T(p, w)$ express that a property p is true in a world w, and $w_1 < w_2$ express that there is a path from w_1 to w_2.

Then the order which prefers a model which changes minimally in one direction is defined as follows. Let M and M' be interpretations with the domain D which are comparable with respect to $< T >$.

$$
\begin{aligned}
&\mathcal{R}(M', M)_\phi \text{ if and only if} \\
&\forall\phi_p{\in}\Phi_D((M' \models_{\phi_p} T(p,0)) \equiv (M \models_{\phi_p} T(p,0)))\wedge \\
&\forall\phi_{w_1}{\in}\Phi_D(\\
&\quad ((M \models_{\phi_{w_1}} 0 < w_1)\wedge \\
&\quad \forall\phi_{w_1 w_2}{\in}\Phi_D((M \models_{\phi_{w_1 w_2}} (0 < w_2 \wedge w_2 < w_1)) \supset \\
&\qquad \forall\phi_{w_1 w_2 p}{\in}\Phi_D((M \models_{\phi_{w_1 w_2 p}} T(p,w_2)) \equiv (M' \models_{\phi_{w_1 w_2 p}} T(p,w_2))))) \supset \\
&\quad (\forall\phi_{w_1 p}{\in}\Phi_D((M' \models_{\phi_{w_1 p}} T(p, last(w_1)) \not\equiv T(p,w_1)) \supset \\
&\qquad\qquad\qquad\qquad (M \models_{\phi_{w_1 p}} T(p, last(w_1)) \not\equiv T(p,w_1)))\wedge \\
&\qquad \exists\phi_{w_1 p}{\in}\Phi_D((M \models_{\phi_{w_1 p}} T(p, last(w_1)) \not\equiv T(p,w_1))\wedge \\
&\qquad\qquad\qquad\qquad (M' \models_{\phi_{w_1 p}} T(p, last(w_1)) \equiv T(p,w_1))))).
\end{aligned}
$$

This definition means informally that for every node w, if M and M' agree on the interpretation of T from root to $last(w)$, then M' changes strictly less than M at the point from $last(w)$ to w.

We call a minimal models of the above wdr *minimal change models* because in the minimal models, the change of the property is minimized in the direction from root to leaf. Let τ be similar to T. Then the syntactic definition of the minimal change models, $A(T) \wedge \neg\exists\tau(A(\tau) \wedge R(\tau))$, is as follows.

$$
\begin{aligned}
&A(T) \wedge \neg\exists\tau(A(\tau)\wedge \\
&\quad \forall p(\tau(p,0) \equiv T(p,0))\wedge \\
&\quad \forall w_1(\\
&\qquad (0 < w_1\wedge \\
&\qquad \forall w_2((0 < w_2 \wedge w_2 < w_1) \supset \\
&\qquad\quad \forall p(T(p,w_2) \equiv \tau(p,w_2)))) \supset \\
&\qquad (\forall p((\tau(p, last(w_1)) \not\equiv \tau(p,w_1)) \supset (T(p, last(w_1)) \not\equiv T(p,w_1)))\wedge \\
&\qquad\quad \exists p((T(p, last(w_1)) \not\equiv T(p,w_1)) \wedge (\tau(p, last(w_1)) \equiv \tau(p,w_1)))))).
\end{aligned}
$$

We show an example of derivation in the following linear-structured inheritance system.

Animals do not normally fly.
Mammals are animals.
Bats are mammals and normally fly.

We can express the above information as the following axioms. Let $A(T)$ be:

$$
\begin{aligned}
&\neg T(fly, animal) \wedge T(fly, bat)\wedge \\
&\forall p(p = fly) \wedge \forall w(w = animal \vee w = mammal \vee w = bat)\wedge \\
&0 = animal \wedge last(mammal) = animal \wedge last(bat) = mammal\wedge \\
&animal \neq mammal \wedge mammal \neq bat \wedge animal \neq bat\wedge \\
&\forall w\forall w'(w < w' \equiv (\\
&\quad (w = animal \wedge w' = mammal)\vee \\
&\quad (w = mammal \wedge w' = bat)\vee \\
&\quad (w = animal \wedge w' = bat))).
\end{aligned}
$$

Let $\tau = \lambda p\lambda w(p = fly \wedge w = bat)$. We consider $\neg\exists\tau(A(\tau) \wedge R(\tau))$. $A(\tau)$ and $\forall p(\tau(p,0) \equiv T(p,0))$ are true if we assume $A(T)$. Therefore $R(\tau)$ becomes:

$$\neg((T(fly, mammal) = \tau(fly, mammal)) \supset$$
$$(((\tau(fly, animal) \not\equiv \tau(fly, mammal)) \supset$$
$$(T(fly, animal) \not\equiv T(fly, mammal)))\wedge$$
$$((T(fly, animal) \not\equiv T(fly, mammal))\wedge$$
$$(\tau(fly, animal) \equiv \tau(fly, mammal)))))$$

It becomes:

$$\neg T(fly, mammal).$$

Therefore in the minimal change models, mammals do not normally fly.

5.6 Relative Plausibility

We have presented a formalism of relative plausibility in [Satoh88]. The relative plausibility expresses that a certain wff $E(\mathbf{P}, x_1, ..., x_n)$ is more plausible than another wff $E'(\mathbf{P}, x_1, ..., x_n)$ where $\mathbf{P}$ is a tuple of all predicate constants occurring in wffs E and E', and $x_1, ..., x_n$ are all free individual variables occurring in wffs.

The order of models defined as follows. (Here, We present an order in the case that only one information of relative plausibility is known. A general case is found in [Satoh88].)

Let M and M' be interpretations with the domain D which are comparable with respect to $\mathbf{P}$.

$\mathcal{R}(M', M)$ if and only if

$$\forall\phi_{x_1}\in\Phi_D...\forall\phi_{x_1x_2...x_n}\in\Phi_D($$
$$((M \models_{\phi_{x_1...x_n}} \neg E(\mathbf{P}, x_1, ..., x_n)) \wedge (M \models_{\phi_{x_1...x_n}} E'(\mathbf{P}, x_1, ..., x_n))) \supset$$
$$((M' \models_{\phi_{x_1...x_n}} \neg E(\mathbf{P}, x_1, ..., x_n)) \vee (M' \models_{\phi_{x_1...x_n}} E'(\mathbf{P}, x_1, ..., x_n))))\wedge$$
$$\neg\forall\phi_{x_1}\in\Phi_D...\forall\phi_{x_1x_2...x_n}\in\Phi_D($$
$$((M' \models_{\phi_{x_1...x_n}} \neg E(\mathbf{P}, x_1, ..., x_n)) \wedge (M' \models_{\phi_{x_1...x_n}} E'(\mathbf{P}, x_1, ..., x_n))) \supset$$
$$((M \models_{\phi_{x_1...x_n}} \neg E(\mathbf{P}, x_1, ..., x_n)) \vee (M \models_{\phi_{x_1...x_n}} E'(\mathbf{P}, x_1, ..., x_n)))).$$

The syntactic definition of the relative plausibility is as follows. Let $\mathbf{p} =< p_1, ..., p_n >$ be similar to $\mathbf{P}$.

$$A(\mathbf{P}) \wedge \neg\exists p_1...\exists p_n(A(\mathbf{p})\wedge$$
$$\forall x_1...\forall x_n((\neg E(\mathbf{P}, x_1, ..., x_n) \wedge E'(\mathbf{P}, x_1, ..., x_n)) \supset$$
$$(\neg E(\mathbf{p}, x_1, ..., x_n) \vee E'(\mathbf{p}, x_1, ..., x_n)))\wedge$$
$$\forall x_1...\forall x_n((\neg E(\mathbf{p}, x_1, ..., x_n) \wedge E'(\mathbf{p}, x_1, ..., x_n)) \supset$$
$$(\neg E(\mathbf{P}, x_1, ..., x_n) \vee E'(\mathbf{P}, x_1, ..., x_n))).$$

We show an example of derivation in the propositional case. Let S_1 be a symptom, and D_1 and D_2 be diseases. And we have the following information.

1. A patient suffers from D_1 or D_2.

2. S_1 is found.

3. If S_1 is found then a patient suffers from D_1 more likely than D_2.

Then we express the above information as follows.

$A(S_1, D_1, D_2) = (D_1 \vee D_2) \wedge S_1$.
$S_1 \supset D_1$ is more plausible than $S_1 \supset D_2$.

Then we can make the following derivation.

$$(D_1 \vee D_2) \wedge S_1 \wedge \neg\exists s_1 \exists d_1 \exists d_2((d_1 \vee d_2) \wedge s_1 \wedge$$
$$((\neg(S_1 \supset D_2) \wedge (S_1 \supset D_1)) \supset (\neg(s_1 \supset d_2) \vee (s_1 \supset d_1))) \wedge$$
$$\neg((\neg(s_1 \supset d_2) \wedge (s_1 \supset d_1)) \supset (\neg(S_1 \supset D_2) \vee (S_1 \supset D_1)))).$$

Suppose $s_1 = S_1$ and $d_1 = S_1$ and $d_2 = \neg S_1$. Then we can derive the following from the above wff.

$$(D_1 \vee D_2) \wedge S_1 \wedge \neg((S_1 \vee \neg S_1) \wedge S_1 \wedge$$
$$((\neg(S_1 \supset D_2) \wedge (S_1 \supset D_1)) \supset (\neg(S_1 \supset \neg S_1) \vee (S_1 \supset S_1))) \wedge$$
$$\neg((\neg(S_1 \supset \neg S_1) \wedge (S_1 \supset S_1)) \supset (\neg(S_1 \supset D_2) \vee (S_1 \supset D_1)))).$$

It is reduced to:

$$(D_1 \vee D_2) \wedge S_1 \wedge \neg(\neg(\neg D_2 \vee D_1)),$$

which is equivalent to:

$$D_1 \wedge S_1.$$

Therefore, we can conclude that a patient suffers from a disease D_1 from the above information of relative plausibility.

6 Discussion

In this section, we discuss two limitations of our framework in reasoning of inequality and conditional probability.

6.1 Reasoning of Inequality

Etherington[Etherington87] shows that Lifschitz's general circumscription without any variable terms cannot derive $a \neq b$ from $\mathbf{T}$ (no proper axioms).

Arima[Arima88] formalizes the limitation of circumscription in reasoning inequality and shows that it is impossible to infer $a \neq b$ from $\mathbf{T}$ even with variable terms.

Similar problem can also arise in our framework [2]. It is because in our framework, preference only can be defined over the *comparable* interpretations which are different each other only in the interpretation of some predicates.

[2] Actually, in our framework, if we gives the following preference order, we can infer $a \neq b$ from $\mathbf{T}$. This can be regarded as a difference between Lifschitz's general circumscription and our framework. However, it is not applicable to all the cases because this solution is in danger of contradiction.

$$\mathcal{R}(M', M) \text{ if and only if } M \models a = b.$$

6.2 Conditional Probability

In the conditional probability, if more information is obtained, probabilities for some formulas may be changed. This corresponds to the change of preference order in our framework if models are changed. However, since preference order in our framework is fixed over interpretations, we cannot change it even if a set of models are changed.

In default logic, this problem is half solved thanks to *prerequisites* of *defaults*. Since defaults cannot apply if prerequisites are not derived, defaults with prerequisites can add the extra preference order when the formula in prerequisite is known to be true. And this is the reason why default with prerequisite can not translate into circumscription as stated in [Imielinski87]. However, even in default logic, we cannot express the deletion of the preference. Therefore, we need other framework to express conditional probability precisely.

Acknowledgments

I would like to thank Jun Arima at ICOT for helpful discussions. My special thanks must go to Yukihide Takayama who is open-minded enough to read earlier version of the paper and give instructive comments.

References

[Arima88] Arima, J., The anonym problem: a weak point of circumscription on equality, *ICOT-TR400* (1988).

[Etherington86] Etherington, D.W., Formalizing nonmonotonic reasoning systems, *Artificial Intelligence* **31** (1987) 41–85.

[Etherington87] Etherington, D.W., Relating default logic and circumscription, *Proceedings IJCAI-87* (1987) 495–498.

[Hanks87] Hanks, S., and McDermott, D., Nonmonotonic logic and temporal projection, *Artificial Intelligence* **33** (1987) 379–412.

[Imielinski87] Imielinski, T., Results on translating defaults to circumscription, *Artificial Intelligence* **32** (1987) 131–146.

[Kautz86] Kautz, H., The logic of persistence, in: *Proceedings AAAI-86* (1986) 401–405.

[Lifschitz84] Lifschitz, V., Some results on circumscription, in: *Proceedings AAAI Workshop on Nonmonotonoic Reasoning* (1984) 151–164.

[Lifschitz86] Lifschitz, V., Pointwise circumscription: preliminary report, in: *Proceedings AAAI-86* (1986) 406–410.

[McCarthy80] McCarthy, J., Circumscription — a form of non-monotonic reasoning, *Artificial Intelligence* **13** (1980) 27–39.

[McCarthy86] McCarthy, J., Applications of circumscription to formalize commonsense knowledge, *Artificial Intelligence* **28** (1986) 89–116.

[McDermott80] McDermott, D., and Doyle, J., Nonmonotonic logic I, *Artificial Intelligence* **13** (1980) 41–72.

[Reiter80] Reiter, R., A logic for default reasoning, *Artificial Intelligence* **13** (1980) 81–132.

[Satoh87] Satoh, K., Minimal change - a criterion for choosing between competing models -, *ICOT-TR316*, ICOT (1987), also in: *Proceedings International Workshop on Machine Learning, Meta-Reasoning, and Logics* (1988) 195–221.

[Satoh88] Satoh, K., Relative plausibility based on model ordering: preliminary report, *ICOT-TR429*, ICOT (1987).

[Shoham86] Shoham, Y., Chronological ignorance: time, nonmonotonicity, necessity and causal theories, in: *Proceedings AAAI-86* (1986) 389–393.

[Shoham87] Shoham, Y., Nonmonotonic logics: meaning and utility, in: *Proceedings IJCAI-87* (1987) 388–393.