TR-357

On the Semantics of Hypothetical Reasoning and Truth Maintenance

by

K. Inoue

March, 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

On the Semantics of Hypothetical Reasoning and Truth Maintenance

Katsumi Inoue

ICOT Research Center,
1–4–28, Mita, Minato-ku, Tokyo 108, Japan
phone: +81-3-456-3192
telex: ICOT J 32964
csnet: inoue%icot.jp@relay.cs.net
uucp: {kddlab,mit-eddie,ukc,enea,inria}!icot!inoue

ICOT Technical Report TR-356
March 1988

Abstract

This paper describes a logical framework for hypothetical reasoning. Hypothetical reasoning is a kind of non-monotonic reasoning, which is desirable to have when dealing with incomplete knowledge in problem solving, making hypotheses for this type of knowledge, with which we can infer the kinds of formula that hold based on the different kinds of hypothesis. Two different aspects of hypothetical reasoning, which are abductive reasoning and consistency maintenance, are formalized in the unified model theory. The paper also shows that the semantics for default logic can be partly incorporated in the model theory. In this simple and clear logical framework for hypothetical reasoning, truth maintenance systems that have been widely used, but lack model-theoretic semantics, can be analyzed theoretically. While all models of sets of beliefs are maintained by de Kleer's ATMS, only one model of some set of beliefs is selected by Doyle's TMS. These results show that hypothetical reasoning is an important subcase of default logic for which efficient theorem proving techniques exist.

Paper Length: 4000 words.

Topic: Automated Reasoning.

Keywords: Truth Maintenance, Abductive Reasoning, Default Logic, Hypothetical Reasoning, Model Theory.

1. Introduction

In the real world, we are often forced to draw some conclusion even if complete information is not available in problem solving. This inference may have to anticipate the possibility of later revisions of beliefs. An approach to this kind of *non-monotonic reasoning* that has been pursued to use in practice is seen in the field of *truth maintenance systems* (TMSs) such as [Doyle 79] and [de Kleer 86]. Although TMSs have been widely used as non-monotonic reasoning systems, there has been little formal research on them. Because they lack clear semantics, TMSs have not been well defined or understood in comparison with non-monotonic logics.

The motivation of this research was to formalize reasoning systems with incomplete knowledge, that is, to clarify model-theoretic semantics of such reasoning systems. One example of this research can be seen in the *clause management system* (CMS) [Reiter & de Kleer 87], which simply formalizes de Kleer's *assumption-based TMS* (ATMS). We propose an alternative approach for reasoning with incomplete knowledge, in terms of *hypothetical reasoning*, which is more general than the CMS. Hypothetical reasoning is desirable when dealing with incomplete knowledge in problem solving, and it makes hypotheses for this type of knowledge, with which we can infer what formulas hold based on what hypotheses. In this paper, firstly, two different aspects of hypothetical reasoning are shown, and are then represented in a simple logical framework, where Reiter's *default logic* [Reiter 80] is shown to be incorporated. Secondly, the model theory for the ATMS and Doyle's TMS are presented, based on this logical framework.

2. Utility of Hypothetical Reasoning

When we *know* a formula, $p \supset q$, we can say that if p is *assumable*, q holds under *assumption* p by *hypothetical syllogism* in traditional logic. In hypothetical reasoning, these assumptions (or *hypotheses*) given by a problem solver are not guaranteed to be always true. Typical uses of hypothetical reasoning are broadly classified as follows.

1. Consistency maintenance in knowledge bases
In constraint satisfaction, when we select from alternatives but there is not enough information to select one, we assume one and process reasoning further from there. Because assumptions and formulas derived from them are not guaranteed to be true, if a contradiction occurs in the reasoning process, we must remove the original assumptions and select other ones instead. This type of reasoning is sometimes called *assumption-based reasoning*. Typical AI systems with this mechanism are implemented as TMSs, which maintain the *consistency* of knowledge bases, so that they can deal with contradictions.

2. Abductive reasoning
In model-based reasoning, when an observation of a system's behavior is found, we want to know *hypotheses* that explain the observation with knowledge of descriptions of the system and devices. This type of reasoning is called *abductive reasoning*, and in engineering, can be applied directly to diagnostic reasoning such as the theory formation in [Poole 86]. When the specifications are given as observations, it can also be applied to design such as the resolution residue in [Finger & Genesereth 85].

Formal theories are studied mainly in abductive reasoning, such as [Poole 86] and [Reiter & de Kleer 87]. However, consistency maintenance mechanisms are pursued algorithmically rather than theoretically. We suggest that both types of hypothetical reasoning should be understood in the same model theory, that is, the latter case can be analyzed in the same way as the former. These two types of hypothetical reasoning simply correspond to different aspects of the same logical framework.

3. General Formalism

This section considers a model theory for hypothetical reasoning. First, reasoning from knowledge is formalized. Then we show a logical framework for supporting hypotheses which sanction given propositional well-formed formulas. The framework corresponds to a model theory for abductive reasoning. Next, reasoning from beliefs is formalized within the framework. This formalism reflects the property of consistency maintenance and it is compared with Reiter's default logic.

3.1 Supporting Hypotheses

Let U be a set of finite propositional symbols, and L be a propositional language with U and logical connectives, and the set of *well-formed formulas* (*wffs*). [1] An *interpretation*, I, of U is defined as an element of 2^U such that for each $\alpha \in I$, α is supposed to be assigned to true. The relation, "$\models$" ($\subseteq 2^U \times L$), can be defined in the usual way. Let $M \in 2^U$ be an interpretation and W be a set of wffs in L, then the fact that M *satisfies* W is defined as $M \models \alpha$ for all $\alpha \in W$. In this case, we say that M is a *model* of W. W is said to be *satisfiable* if it has at least one model. The set of all models of W is denoted as $MOD(W)$. We say that for a wff, $\alpha \in L$, W *entails* α (written $W \models \alpha$), if for each model, $M \in MOD(W)$, $M \models \alpha$ holds. A *theory* of W (denoted $Th(W)$) is a set of wffs closed under entailment. Finally, we say that $M \in MOD(W)$ is a *minimal model* of W iff $M' \in MOD(W)$ and $M' \subseteq M$ only if $M' = M$.

Definition 1. A set of *premises*, Σ, is a satisfiable set of wffs. A *set of knowledge of* Σ is defined as $Th(\Sigma)$. A wff, $w \in L$, is *knowledge of* Σ iff $\Sigma \models w$. □

Definition 2. A wff, $d \in L$, is *indefinite with respect to* Σ iff $\Sigma \not\models d$ and $\Sigma \not\models \neg d$. A set of all indefinite wffs with respect to Σ is denoted as $DMAX_{(\Sigma)}$. [2] A set of wffs, $D_{(\Sigma)}$, is a set of *hypotheses with respect to* Σ iff $D_{(\Sigma)}$ is a subset of $DMAX_{(\Sigma)}$. A set of wffs, $E_{D_{(\Sigma)}}$, is an *environment of* $D_{(\Sigma)}$ iff $E_{D_{(\Sigma)}} \in 2^{D_{(\Sigma)}}$. We denote a pair of Σ and $E_{D_{(\Sigma)}}$ as $(\Sigma, E_{D_{(\Sigma)}})$. In the subsequent discussion, we omit the subscript "$_{(\Sigma)}$" in D or $DMAX$ when they are clear in the situation, and we simply say "indefinite wff" or "set of hypotheses", omitting "with respect to Σ". □

[1] While we use the propositional language, L, to make the discussion clear, all concepts in this paper can be extended to have a subset of the first order predicate calculus (FOPC), where it is assumed to be function-free and each formula is assumed to be universally quantified. The set of all ground atomic formulas in FOPC corresponds to U defined above.

[2] The indefiniteness of Definition 2 is more general than that of [Minker 82], where the indefiniteness is determined not by all models but by *minimal* models.

Definition 3. Let w be a wff in L. A wff, $d \in D$, is a *supporting hypothesis for w with respect to* (Σ, D) iff

$$\Sigma \models d \supset w$$

holds. The set of all supporting hypotheses for w with respect to (Σ, D) is denoted as $SMAX(\Sigma, D, w)$. $\square$

Proposition 1. Suppose that $d \in SMAX(\Sigma, D, w)$. Then, (1) $\Sigma \cup \{d\}$ is satisfiable, and (2) $\Sigma \cup \{d\} \models w$. $\square$

Proposition 2. (1) If $\Sigma \models w$, then for any $D \in 2^{DMAX}$, $SMAX(\Sigma, D, w) = D$.
(2) $w \in D$ iff $w \in SMAX(\Sigma, D, w)$.
(3) $\Sigma \models \neg w$ iff $SMAX(\Sigma, DMAX, w) = \phi$. $\square$

Intuitively, $d \in SMAX(\Sigma, D, w)$ is a supplementary wff with which w is entailed by Σ, keeping consistency with Σ. Given w, the computation of a set of supporting hypotheses for w with respect to (Σ, D) corresponds to abductive reasoning. [3] We should pay attention to the *principle of parsimony*, that is, such supporting hypotheses are to be minimal as follows.

Definition 4. Let w be a wff in L. A wff, $d \in D$, is a *minimal supporting hypothesis for w with respect to* (Σ, D) iff

$$d \in SMAX(\Sigma, D, w) \;\wedge\; \neg\exists d' \in SMAX(\Sigma, D, w).\, (d \not\equiv d' \;\wedge\; d \supset d')\,.$$

The set of all minimal supporting hypotheses for w with respect to (Σ, D) is denoted as $SMIN(\Sigma, D, w)$. If $\Sigma \models w$, then $SMIN(\Sigma, D, w)$ is denoted as $\top$. $\square$

Note that for each $d \in SMIN(\Sigma, D, w)$, $MOD(\{d\})$ is maximal (in terms of set inclusion) in $SMAX(\Sigma, D, w)$. To analyze the properties of $SMIN(\Sigma, D, w)$, we define the set of all models of Σ that satisfy at least one hypothesis in an environment, E_D, as follows.

$$M(\Sigma, E_D) \stackrel{\text{def}}{=} \bigcup_{e \in E_D} MOD(\Sigma \cup \{e\})\,.$$

Proposition 3. (1) $M(\Sigma, SMIN(\Sigma, D, w)) \subseteq MOD(\Sigma \cup \{w\})$. (2) If $w \in D$, then $M(\Sigma, SMIN(\Sigma, D, w)) = MOD(\Sigma \cup \{w\}) = MOD(\Sigma) \cap MOD(\{w\})$. $\square$

Theorem 4. $M(\Sigma, SMIN(\Sigma, D, w)) = M(\Sigma, SMAX(\Sigma, D, w))$. $\square$

Theorem 4 shows that only from $SMIN(\Sigma, D, w)$, all models of Σ that contain at least one hypothesis in $SMAX(\Sigma, D, w)$ can be computed.

---

[3] In [Poole 86] and [Reiter & de Kleer 87], the requirement for each hypothesis, d, is only its consistency with Σ, i.e., $\Sigma \not\models \neg d$. Our definition with the indefiniteness is more restrictive. The difference arises when d is knowledge of Σ. It is of course not problematic that some knowledge is contained in D. However, the aim of abductive reasoning is to obtain hypotheses that explains an observation when it cannot be derived from Σ. Therefore, indefiniteness is essential in hypothetical reasoning.

Theorem 5 (monotonicity of supporting hypotheses). Let D and D' be two sets of hypotheses. If $D \subseteq D'$, then $M(\Sigma, SMIN(\Sigma, D, w)) \subseteq M(\Sigma, SMIN(\Sigma, D', w))$. □

Example 1. Suppose that Σ is a set of premises, and that D_1 and D_2 are two sets of hypotheses, as follows.

$\Sigma = \{a \supset c,\ b \supset c,\ a \wedge b \supset g,\ \neg g\}$
$D_1 = \{a,\ b\}$
$D_2 = \{a,\ b,\ a \vee b,\ a \supset b,\ c,\ b \wedge c\}$

From this, the following sets of supporting hypotheses can be obtained.

$SMAX(\Sigma, D_1, c) = SMIN(\Sigma, D_1, c) = \{a,\ b\}$
$SMAX(\Sigma, D_2, c) = \{a,\ b,\ a \vee b,\ c,\ b \wedge c\}$
$SMIN(\Sigma, D_2, c) = \{a \vee b,\ c\}$
$SMAX(\Sigma, D_2, b \wedge c) = \{b,\ b \wedge c\}$
$SMIN(\Sigma, D_2, b \wedge c) = \{b\}$ □

Remarks (on the comparison with the CMS). The above definition of a *supporting hypothesis*, s, can be roughly compared with the notion of a 'support', $\neg s$, of the CMS in [Reiter & de Kleer 87]. In the CMS, however, all wffs have to be translated to the clausal normal form and some or all 'prime implicants' of a set of clauses, Σ_C, are considered so that $SMIN(\Sigma_C, DMAX_C, c)$ for any clause c is computed, where $DMAX_C$ is the set of all indefinite clauses with respect to Σ_C. We focus only on model-theoretic semantics, and our formalism with any wffs is more general than one with the clausal normal form. For example, our formalism naturally provides minimal supporting hypotheses for conjunctive observations such as $SMIN(\Sigma, D_2, b \wedge c)$ in Example 1. Moreover, we deal with indefinite wffs not by $DMAX$ but by the *set of hypotheses*, D, which is all that is needed to construct a set of supporting hypotheses for any wff. Since D is given in our formalism, the sets of *beliefs* can be constructed from Σ and environments of D, and they can be characterized in terms of *extensions* in default logic in Section 3.2. □

3.2 Beliefs and Extensions

One technique for *efficient* computation of $SMIN(\Sigma, D, w)$ for a wff, $w \in L$, can be considered to be a mechanism to keep dependencies or the original assumptions with w in TMSs, as stated in Section 4. Therefore, the above formalism for abductive reasoning can also be a basis for a model theory of consistency maintenance, as follows.

Definition 5. Suppose that E_D is an environment of D such that $\Sigma \cup E_D$ is satisfiable. A *set of beliefs of* (Σ, E_D) (denoted $B(\Sigma, E_D)$) is defined as $Th(\Sigma \cup E_D)$. A wff, $w \in L$, is a *belief of* (Σ, E_D) iff $w \in B(\Sigma, E_D)$. □

Theorem 6. Let w be a wff in L, then,
(1) if w is a belief of (Σ, E_D) and $d \in D$ where $d \supset d'$ and $d' = \bigwedge_{e \in E_D} e$, d is a supporting hypothesis for w with respect to (Σ, D), and
(2) if d is a supporting hypothesis for w with respect to (Σ, D), then for each environment E_D of D where E_D contains d' such that $d' \supset d$, w is a belief of (Σ, E_D). □

Theorem 6 gives a bridge between consistency maintenance and abductive reasoning. Theorem 7 shows that $MOD(B(\Sigma, E_D))$ decreases monotonically as E_D increases.

Theorem 7 (model monotonicity of hypotheses). Let $d \in D$, and E, E' be two environments of D. Suppose that $\Sigma \cup E$, $\Sigma \cup E \cup \{d\}$ and $\Sigma \cup E'$ are satisfiable.
(1) If $d \notin E$, then $MOD(B(\Sigma, E \cup \{d\})) = MOD(B(\Sigma, E)) - MOD(\{\neg d\})$.
(2) If $E \subseteq E'$, then $MOD(B(\Sigma, E')) \subseteq MOD(B(\Sigma, E))$. □

Remark. The monotonicity with respect to premises does not hold. Suppose that $\Sigma \cup E_D$ is satisfiable. When a new premise p is added to Σ, $\Sigma \cup \{p\} \cup E_D$ might not be satisfiable. Then, we should discard some hypotheses from E_D to obtain satisfiability of the new set of beliefs, which might not have fewer models than the previous one. □

A set of hypotheses, D, is very closely related to a restricted case of a set of *normal defaults* in default logic [Reiter 80]. For each $d \in D$, d corresponds to the consequent of a normal default of the form $:M\, d \,/\, d$. We now define a fixed point operator NM as follows:

$$NM_{(\Sigma,D)}(S) \stackrel{\text{def}}{=} B(\Sigma, \{\, d \mid d \in D,\ \neg d \notin S\}).$$

Definition 6. An *extension of* (Σ, D) is the fixed point of $NM_{(\Sigma,D)}$, that is, S is an extension of (Σ, D) iff $NM_{(\Sigma,D)}(S) = S$. □

All results of normal default theories in default logic are applicable to our formalism. In particular, the property of *semi-monotonicity* corresponds to the result of Theorem 7. Let E'_D be a maximal environment in 2^D obtained by repeatedly applying each hypothesis as long as the set of beliefs is satisfiable, then $B(\Sigma, E'_D)$ is an extension of (Σ, D), where $MOD(B(\Sigma, E'_D))$ is a minimal set of models in $2^{MOD(\Sigma)}$. This result corresponds to the semantics for normal default theories in [Etherington 87].

Example 2 (closed world assumption). The *closed world assumption* (CWA) in [Reiter 78] says that if Σ does not entail an atomic formula, p, then $\neg p$ can be assumed to be true. In our model theory, the CWA can be characterized as:

$$D_{CWA} \stackrel{\text{def}}{=} \{\neg p \mid p \in U \text{ and } p \in DMAX_{(\Sigma)}\}.$$

Then, the set of all models of all extensions of (Σ, D_{CWA}) is equivalent to the set of all minimal models of Σ. This result is an alternative form of the *generalized closed world assumption* (GCWA) in [Minker 82]. □

4. Truth Maintenance Systems

The logical framework for hypothetical reasoning subsumes various TMSs. This section shows how the general formalism is related to TMSs, especially to de Kleer's *assumption-based TMS* (ATMS) [de Kleer 86] and Doyle's non-monotonic *justification-based TMS* (JTMS) [Doyle 79], which are two different kinds of representative TMSs. In the following discussions, we distinguish the ATMS from the JTMS mainly by how the knowledge base is maintained in terms of the model theory. While all models of all extensions of (Σ, D) are maintained by the ATMS, the JTMS selects only the current model of some set of beliefs of (Σ, E_D) reflecting some 'intended meaning'.

Definition 7. A *node* is defined to be associated with each atomic formula in U. An *assumption* is an indefinite atomic formula with respect to Σ. We denote the set of all assumptions as $AMAX$. A *justification* is a wff of the form:

$$\alpha_1 \wedge \ldots \wedge \alpha_m \wedge \neg\beta_1 \wedge \ldots \wedge \neg\beta_n \supset w,$$

where $m, n \geq 0$, and α_i, β_j (called *antecedents*) denote atomic formulas in U, and w (called *consequent*) denotes an atomic formula in U or $\perp$ (*falsity*). Note that any clause can be transformed to a justification. A justification is called *monotonic* iff it contains no negative antecedents (i.e., $n = 0$). A monotonic justification whose consequent is $\perp$ is called *nogood*. A justification is called *non-monotonic* iff it contains at least one negative antecedent (i.e., $n \geq 1$). A set of premises, Σ, is given as a satisfiable set of justifications. We allow for multiple justifications for one single node, w. □

4.1 de Kleer's ATMS: Keeping All Extensions

The ATMS maintains a concurrent representation of all sets of beliefs by labeling each atomic formula with all minimal supporting hypotheses. In other words, the ATMS keeps all models of all extensions of (Σ, D). In the ATMS, Σ is given as a set of monotonic justifications, and a *set of assumptions*, AS, is given as a subset of $AMAX$. Then, in our general model theory, the set of hypotheses, D, is defined as follows.

$$D \stackrel{\text{def}}{=} \{d \mid d = \bigwedge_{a \in A} a,\ A \in 2^{AS},\ \text{and}\ d \in DMAX_{(\Sigma)}\}.$$

For each node w, $SMIN(\Sigma, D, w)$ (called the *label* of w) is maintained. A set of all nodes in a set of beliefs of (Σ, A) $(A \in 2^{AS})$ is called a *context* in the ATMS.

We can generalize the above characterization of the ATMS with our formalism. For example, Σ can be extended to have any non-monotonic justifications, any clauses, or even any wffs. In particular, the generalization of Σ from Horn clauses to general clauses is equivalent to one with the CMS [Reiter & de Kleer 87]. The following results give a characterization of the generalized ATMS.

Theorem 8. Suppose that Σ, AS and D are the same as the above definition, except that Σ is given as a set of any wffs. Let w be a wff in L.
(1) The set of all extensions of (Σ, D) is equivalent to the set of all extensions of (Σ, AS).
(2) $SMIN(\Sigma, D, w) \neq \phi$ iff w is a belief of some extension of (Σ, AS).
(3) $SMIN(\Sigma, D, w) = \top$ iff w is knowledge of Σ. □

In the ATMS, the computation of all extensions of (Σ, AS) is called *interpretation construction*. By the property of *semi-monotonicity* of Theorem 7, each hypothesis can be treated independently. Therefore, interpretation construction can be done simply by applying one assumption after another and ignoring some assumptions, if the ultimate set of all premises, Σ, is explicitly available. However, Σ is usually defined implicitly. Because of the global property of Σ, if Σ grows, then the environment taken into account decreases. To avoid redundant computing during interpretation construction, *dependency-directed search* (DDS) is utilized with the ATMS.

Example 3. Suppose that Σ_1 and Σ_2 are two sets of premises, and that AS_1 and AS_2 are two sets of assumptions. Let D_1 and D_2 be two sets of hypotheses for AS_1 and AS_2, respectively, given by the previous definition. From

$\Sigma_1 = \{\, a \vee b,\ a \wedge c \supset \perp,\ c \supset g \,\}$ and
$AS_1 = \{\, a,\ b,\ c \,\}$,

we can obtain $SMIN(\Sigma_1, D_1, g) = \{c\}$. In this case, we have two extensions of (Σ_1, AS_1), i.e., $B(\Sigma_1, \{a, b\})$ and $B(\Sigma_1, \{b, c\})$. Then, the following premise is added to Σ_1, and Σ_2 is created as

$\Sigma_2 = \Sigma_1 \cup \{\, b \wedge c \supset \perp \,\}$.

Since $\Sigma_2 \models \neg c$, c is no longer indefinite. Therefore, from

$AS_2 = AS_1 - \{\, c \,\}$,

we now obtain $SMIN(\Sigma_2, D_2, g) = \emptyset$. There is only one extension of (Σ_2, AS_2), i.e., $B(\Sigma_2, \{a, b\})$. □

4.2 Doyle's TMS: Selecting One Model

The JTMS focuses on only one model of a set of beliefs at a time. In our model theory, it is implicitly assumed that there are all models of every possible set of beliefs. Therefore, the JTMS is interpreted as an extended reasoning module for the logical framework for hypothetical reasoning, whose task can be characterized as to determine one *current* model, $\gamma_c \in MOD(\Sigma)$, where Σ can contain non-monotonic justifications. Then, (1) a node w is 'believed' (or is labeled *in*) iff $w \in \gamma_c$, and (2) a node w is not 'believed' (or is labeled *out*) iff $w \notin \gamma_c$. The principles to determine γ_c reflect the following 'intended meaning' of each justification.

(1) γ_c is *causal*: For each *in* node, w, there exists a justification, $j \in \Sigma$, whose consequent is w, and all antecedents of j are satisfied by γ_c, that is, all nodes of positive antecedents are *in* and all nodes of negative antecedents are *out*;

(2) γ_c is *well-founded*: Intuitively, for each *in* node, w, w does not depend on itself, or syntactically, there exists a non-circular proof of w within Σ.

First, we consider only principle (1), that is, the case that γ_c is a causal model. Although a causal model is not always a minimal model, when Σ is a set of monotonic justifications, γ_c coincides with a minimal model of Σ. However, the minimality is not sufficient for non-monotonic justifications.

Example 4. Suppose that $\Sigma = \{\, a,\ a \wedge \neg b \supset c \,\}$.
$MOD(\Sigma)$ contains three models $M_1 = \{a, b\}$, $M_2 = \{a, c\}$, $M_3 = \{a, b, c\}$. The ATMS allows for all these models, and the GCWA considers two minimal models, M_1 and M_2. The JTMS selects the only causal model, M_2, as γ_c, because b has no valid reason to be 'believed', while c satisfies the causality. [4] □

Even the causality is not enough sufficient for the JTMS to satisfy the 'intended meaning' of the justifications. The causality selects a subset of $MOD(\Sigma)$ more widely than the JTMS notion when there are circular justifications. [5]

[4] In [Morris 87], the *unidirectional* property of the justifications of the JTMS is syntactically shown to solve the multiple extension problems in default logic. From our model-theoretic point of view, this property corresponds to the *causality*.

[5] This was figured out by a personal communication with Michael Reinfrank.

Example 5. Suppose that $\Sigma = \{ a \supset b,\ \neg a \supset c,\ \neg b \supset d,\ a \supset a \}$.
Σ has two causal (and minimal) models, $M_1 = \{a, b\}$ and $M_2 = \{c, d\}$. However, in M_1, the causality of a is obtained by a justification, $a \supset a$, which is tautological and should be ignored. Therefore, only M_2 reflects the 'intended meaning'. □

To analyze the well-foundedness we define a *priority relation with respect to minimality* as follows: for each justification except nogood, (1) negative antecedents tend to be more minimized than consequent, and (2) positive antecedents tend to be no less minimized than consequent. We say that a model is *well-founded* iff it satisfies a non-circular priority relation and that a model is *admissible* iff it is causal and well-founded. In Example 5, as M_2 is well-founded, it is admissible, but M_1 is not.

Remarks. The semantics for admissibility is closely related to the *perfect model* semantics for the *stratified* deductive databases in [Przymusinski 86]. A model M is *preferable* to M' iff for each $\alpha \in M - M'$, there exists $\beta \in M' - M$ such that β is more minimized than α. A model M is said to be *perfect* iff there is no model preferable to M. Note that every perfect model is minimal and that there may be no perfect model of Σ. In Example 5, M_2 is preferable to M_1 and M_2 is perfect. □

Theorem 9. A perfect model of Σ is admissible. □

The truth maintenance process of the JTMS involves finding an admissible model from the set of all models. Whenever new justifications are added to Σ, the process is repeated. If a nogood justification is added to Σ, an admissible model might not be obtained and DDS will be provided with the JTMS. In this case, the JTMS switches the set of beliefs (called *belief revision*). Syntactically, DDS produces some extra justifications and adds them to Σ. From the model-theoretic point of view, belief revision corresponds to choosing some model from the set of models of a new set of beliefs. We need a criterion for belief revision; intuitively, this is done by *minimizing the set difference* between the previous model, γ_c, and the new possible model of Σ.

4.3 Assumption-based versus Justification-based

The construction of the ATMS is more straightforward than the JTMS in the model theory for hypothetical reasoning. The 'brave' character of default logic, where each extension is treated as an acceptable set of beliefs, is very close to the notion of allowing for all models of sets of beliefs in the ATMS. A major problem of the JTMS is that the algorithm and the data structure are too complex, and as a result, the formal analysis is very difficult. Our formalism for the JTMS is much simpler and clearer, and it appears to give one natural way of interpreting the JTMS as an extended reasoning module for the ATMS, so that the *current model* is selected from all models of sets of beliefs.

However, there is a significant difference between the ATMS and the JTMS. While the ATMS can have an explicit set of assumptions, AS, the JTMS implicitly defines $AMAX$ (and/or D_{CWA}). [6] Therefore, the role of a set of hypotheses in the JTMS

[6] If we could specify a set of variable propositions, the resulting JTMS would be very close to the *extended closed world assumption* (ECWA) in [Gelfond *et al.* 86].

is less important than in the ATMS. The spirit of the JTMS lies in the 'intended meaning' (such as the causality and the well-foundedness) of the justifications, so that the expected model can be obtained.

5. Conclusion

This paper presented a logical framework for hypothetical reasoning, which formalizes reasoning systems with incomplete knowledge. The model theory was shown to be the unified formalism for both abductive reasoning and consistency maintenance in knowledge bases. The key idea is that hypotheses are treated as indefinite wffs, making it possible to relate them to restricted cases of normal defaults. The paper also described the formalism for TMSs in our model theory for hypothetical reasoning. These results show that hypothetical reasoning is an important subcase of default logic for which efficient theorem proving techniques exist.

References

[de Kleer 86] de Kleer, J., "An Assumption-based TMS", *Artificial Intelligence* **28** (1986), pp.127–162.

[Doyle 79] Doyle, J., "A Truth Maintenance System", *Artificial Intelligence* **12** (1986), pp.231–272.

[Etherington 87] Etherington, D. W., "A Semantics for Default Logic", *Proc. IJCAI-87* (1987), pp.495–498.

[Finger & Genesereth 85] Finger, J.J. and Genesereth, M.R., *RESIDUE: A Deductive Approach to Design Synthesis*, Technical Report STAN-CS-85-1035, Department of Computer Science, Stanford University, 1985.

[Gelfond *et al.* 86] Gelfond, M., Przymusinska, H. and Przymusinski, T., "The Extended Closed World Assumption and Its Relationship to Parallel Circumscription", *Proc. ACM SIGACT-SIGMOD Symposium on Principles of Database Systems* (1986), pp.133–139.

[Minker 82] Minker, J., "On Indefinite Databases and the Closed World Assumption", *Proc. 6th Conference on Automated Deduction*, Springer-Verlag (1982), pp.292–308.

[Morris 87] Morris, P., "Curing Anomalous Extensions", *Proc. AAAI-87* (1987), pp.437–442.

[Poole 86] Poole, D. L., *Default Reasoning and Diagnosis as Theory Formation*, Technical Report CS-86-08, Department of Computer Science, University of Waterloo, 1986.

[Przymusinski 86] Przymusinski, T. C., "On the Semantics of Stratified Databases", *Workshop on Foundations of Deductive Databases and Logic Programming* (1986), pp.433–443.

[Reiter 78] Reiter, R., "On Closed World Data Bases", In: *Logic and Data Bases* (Gallair, H. and Minker, J., Eds.), Plenum Press, New York (1978), pp.55–76.

[Reiter 80] Reiter, R., "A Logic for Default Reasoning", *Artificial Intelligence* **13** (1980), pp.81–132.

[Reiter & de Kleer 87] Reiter, R. and de Kleer, J., "Foundations of Assumption-based Truth Maintenance Systems: Preliminary Report", *Proc. AAAI-87* (1987), pp.183–188.