ICOT Technical Report: TR-0815

TR-0815

Acceptable Hypotheses: Semantics for Negation by Default

by

N. Iwayama

November, 1992

© 1992, ICOT

ICOT

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

(03)3456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

Acceptable Hypotheses: Semantics for Negation by Default

Noboru IWAYAMA

Institute for New Generation Computer Technology
1-4-28 Mita, Minato-ku, Tokyo 108, Japan
phone: +81-3-3456-2514, fax: +81-3-3456-1618
email: iwayama@icot.or.jp

Abstract

To discuss the semantics of normal logic programs, the principle of negation as abductive hypotheses is considered an excellent guide to capturing the meaning of negation by default. We state a criterion to classify acceptable hypotheses as negation by default based on this principle, and, by examining programs on the basis of the criterion in terms of Przymusinski's minimal models, provide three types of conditions which should be imposed on acceptable hypotheses. The class of conditions based on the criterion is equivalent to the class of three operational semantics defined by Dung, Kakas and others. One of the conditions is equivalent to the semantics proposed by Brogi et al. based on the expandability of Herbrand interpretations.

1 Introduction

In recent works on the semantics of normal logic programs, the principle explained by the following three statements is recognized as being of great importance:

- Regarding normal logic programs as "positive" logic programs.
- Regarding negation by default as abductive "hypothesis".

- Forcing some conditions on the hypothesis so it is recognized as "proper" negation by default.

There are several contributions based on this principle in discussions on the semantics of normal logic programs. They can be classified into the following three types of approaches according to the frameworks in which the approaches discuss conditions of proper negation by default:

Operational Perspectives The main idea of this approach is to give the semantics of a program P through the notion of derivability in the extended program $P \cup \Delta$ (Δ is a set of negative (or abductive) propositions of the form not_A). There are several variants of this approach, specifically those by Eshghi [Eshghi89] (stable models), Dung [Dung91] (preferred extensions), Kakas [Kakas91] (stable theories) and Kakas et al. [Kakas91, Kakas92](acceptability semantics).

Minimal Models Przymusinski [Przymusinski91] formalized the semantics of programs in terms of stationary expansions, namely theories expanded with some default propositions of the form not_A or $\neg not\_A$, and imposed a stationary condition on the theories with respect to the minimal models of the theories.

Expandability of Herbrand Interpretations Brogi et al. considered admissible interpretations of program P as the semantics of P, namely Herbrand models for $P \cup \Delta$ (Δ is a set of negative (or abductive) propositions of the form not_A). They imposed an admissibility condition on the Herbrand models with respect to their expanded interpretations, those are the Herbrand models for $P \cup \Delta'$ ($\Delta' \supset \Delta$).

We may not find any common property of conditions among the three approaches from the appearances of conditions. However, there seems to exist some common property behind the conditions: for example, in the case of complete admissible supported models, three approaches give us the same answers [Brogi91]. The main contributions of this paper are providing the common property behind the three approaches based on the above principle, for more than the special cases such as 'complete' models, and identifying the relationship among the three approaches more clearly.

To accomplish these, we consider an intuitive criterion to classify the proper negative hypothesis as negation by default (Section 2). Then, we examine the small programs on the basis of the intuitive criterion (Section 3).

As a result of the examinations, we achieve the following:

1. We provide three conditions for negation by default in terms of Przymusinski's minimal models on the basis of the intuitive criterion (Section 3).

2. We redefine the three conditions in terms of expandability of interpretations (Section 4) (one of the three redefined conditions is equivalent to Brogis' admissible supported models).

3. We show that the three redefined conditions are equivalent to operational semantics, they are preferential semantics, stable theory and acceptability semantics (Section 5) (the equivalence to preferential semantics is, actually, given by [Brogi91]).

The third achievement is very important, because operational considerations on semantics are supported by model theoretical formalizations. As a result, we can draw comparisons from model theoretical perspectives in the three operational semantics, i.e., preferential semantics is a special case of stable theories, and stable theories is a special case of acceptability semantics.

2 Validity of Negative Hypotheses Application

As stated in the introduction, we would like to investigate the criterion of classifying hypotheses as proper negation by default, regardless of the framework in which we consider the condition of the choice. At first, we consider the validity of application of default in the following statement:

> A set of default Δ is applicable iff there is no unsuitable case for the application of Δ after Δ is applied.

In this statement, a set of default Δ denotes negative hypotheses which should be considered as negation by default in a certain normal logic program.

Since the meaning of "unsuitable" case is not clear in this statement, we must analyze the unsuitable cases more precisely.

An unsuitable case for the application of Δ is a case where there is some set Δ' of defaults which derives a counter to Δ. But Δ' should not include any hypothesis in Δ, because we have already applied Δ. Of course, we consider only the case where Δ by itself does not derive any counter to itself (we call this notion *coherency* to distinguish it from the classical notion of "consistency"). Now, we show an example.

Example 2.1

$$p \leftarrow not\_q$$

$$q \leftarrow not\_r$$

We would like to discuss the applicability of a set $\Delta = \{not\_q\}$ *(Δ is coherent). Because* $\{not\_r\}$ *derives a counter to* Δ, q, *we conclude that* Δ *is not applicable.*

Even though a counter is derived by some default set Δ' after Δ is applied, there seems to exist cases in which there is no relation to the applicability of Δ. Let us consider the following example by Kakas and Mancarella [Kakas91].

Example 2.2

$$p \leftarrow not\_q$$

$$q \leftarrow not\_q$$

We would like to discuss the applicability of a set $\Delta = \{not\_p\}$ *(Δ is coherent). If* $\Delta' = \{not\_q\}$ *is applied, the counter to* Δ, p, *is derived by* Δ'. *However,* $\{not\_q\}$ *is not coherent because* q *is also derived by* Δ'. *Since* Δ' *is not applicable, there is no possibility of that the counter to* Δ *is derived. So we do not need to consider the case of application of* Δ'.

Here, on the basis of the above discussion on unsuitability of default application, we paraphrase the above statement for the validity of default application:

> **[The Statement of Default Applicability]** A set of default Δ is applicable iff the application of any default set Δ' other than Δ does not derive any counter to Δ after Δ is applied, or Δ' is not applicable after Δ is applied.

This statement contains recursiveness because it refers to a case where Δ' is not *applicable*. In the next section, we clarify the not-applicable case to break the recursion in the above statement by examining the programs.

3 Acceptability Supported by Minimal Models

The paraphrased statement in the previous section is still rather vague. In [Przymusinski91], Pryzmusinski provides us with an elegant framework to deal with negation by default as abductive hypotheses. The framework is based on minimal models. In the following, we examine typical normal logic programs in terms of Pryzmusinski's minimal models to formalize the above statement.

We refer to the terminologies of logic programming. A normal logic program P is a set of clauses of the form

$$A \leftarrow L_1, ..., L_n$$

where A is an atom and $L_1, ..., L_n$ ($n \geq 0$) are literals. Negative literals in clause bodies are denoted by *not*A to clearly distinguish between negation by default (or by failure) "*not*" and classical negation "$\neg$". We consider only (possibly infinite) propositional programs[1].

Let P be a program and HB be the Herbrand base associated with P. As mentioned above, negative literals *not*A are dealt with as positive atoms *not_A*, where *not_A* is a new propositional symbol. In other words, P is transformed into its positive version P' by replacing each negative literal *not*A in P's clause bodies with the corresponding positive literal *not_A*. The positive literal *not_A* is not only a propositional symbol but an abductive hypothesis. We denote the set of abductive hypotheses as not_HB ($= \{not\_A | A \in HB\}$).

In the following, we are going to explain Przymusinski's minimal models in [Przymusinski91]. Given a logic program P, an expansion T is a theory obtained by adding some default propositions of the form *not_A* or $\neg$*not_A* to P. Given a theory T, a Herbrand interpretation I of T is identified with a subset of the Herbrand base $HB \cup not\_HB$. Given an interpretation I, a default or objective proposition F of the form not_A or A is true in I if

[1]This way of restriction on programs is well known in literature.

$F \in I$, otherwise it is false. We say that I is a model of T if I satisfies all the statements in T.

Definition 3.1 *A model M of theory T is* minimal *iff there is no model N of T s.t. $N \subset M$ and N coincides with M on default propositions.*

The class of all minimal models of T is denoted as MIN(T). $T \models_{MIN} F$ means that a formula F is true in all minimal models of T, and $T \models_M F$ means that a formula F is true in minimal model M.

Based on the concept of minimal models, we consider the coherency of hypotheses before examining the programs. In the sequel, $\Delta \subseteq not\_HB$. When we assume some hypotheses as negation by default, the hypotheses should be coherent at least. Here, we give the following definition of coherency of the hypotheses, where we use the immediate consequence operator T_P for a positive program P.

Definition 3.2 *A set of hypotheses Δ is* coherent *to a program P iff there exists no proposition A s.t. $A \in T_{P \cup \Delta} \uparrow \omega$ and $not\_A \in T_{P \cup \Delta} \uparrow \omega$.*

This definition is equivalent to the following model theoretical definition.

Definition 3.3 *Coherent Hypotheses*
A set of hypotheses Δ is coherent to the program P iff

$$\exists M \in MIN(T), \forall A \in HB, T \models not\_A \Rightarrow T \models_M \neg A.$$

Theorem 3.1 *Both definitions are equivalent.*

Proof: By proposition 4.1.□

At this point, we are ready to examine the programs in terms of minimal models.

Example 3.1 *Let us consider the following program and a hypothesis $\Delta = \{not\_p\}$ (Δ is coherent) which we would like to apply.*

$$p \leftarrow not\_q,$$

$$q \leftarrow not\_p.$$

Here, we show the minimal models for $T = P \cup \Delta$:

$$\begin{array}{llllll} M_1: & \{not\_p, & not\_q, & p, & q\} \\ M_2: & \{not\_p, & \neg not\_q, & \neg p, & q\} \end{array}$$

On the basis of the statement of default applicability in the previous section, since there is no counter to Δ in M_2, we do not need to consider this case. In M_1, Δ' ($= \{not\_q\}$) is not applicable because Δ' is incoherent ($T \models_{M_1} not\_q$ but $T \models_{MIN} q$).

This example is suggestive of the condition that Δ' is not applicable after Δ is applied. The next definition is based on this examination.

Definition 3.4 *Admissible hypotheses supported by minimal models*
A coherent hypotheses set Δ for a program P ($T = P \cup \Delta$) is admissible acceptable iff T satisfies the condition:

> $\forall N \in MIN(T), \forall A \in HB,$
> $T \models not\_A$ *and* $T \models_N A \Rightarrow \exists not\_A' \notin \Delta, T \models_N not\_A'$ *and*
> $T \models_{MIN} A'$.

Admissible(P) denotes the class of all admissible hypotheses set for a program P.

We show the definition of stationary expansions in [Przymusinski91] to make a comparison with the above definition.

Definition 3.5 *(Stationary expansions)*
A consistent theory $T = P \cup \Delta$, with $\Delta \subseteq not\_HB \cup \neg not\_HB$, is a stationary expansion of a program P iff $\forall A \in HB$, T satisfies the conditions:

1. $T \models not\_A \Leftrightarrow T \models_{MIN} \neg A$.
2. $T \models \neg not\_A \Leftrightarrow T \models A$.

Here, we use the set of $\neg not\_HB = \{\neg not\_A | not\_A \in not\_HB\}$. We say that the admissibility supported by minimal models is a generalization of stationarity as shown in the following proposition.

Proposition 3.1 *For a stationary expansion $P \cup \Delta$, a hypotheses set of $\Delta \cap not\_HB$ is admissible.*

We proceed to the next example.

Example 3.2 *Let us consider the following program and a hypothesis* $\Delta = \{not\_p\}$ *(Δ is coherent) which we would like to apply.*

$$p \leftarrow not\_q,$$

$$q \leftarrow not\_q.$$

Here we show the minimal models for $T = P \cup \Delta$

$$\begin{array}{llll} M_1: & \{not\_p, & not\_q, & p, & q\} \\ M_2: & \{not\_p, & \neg not\_q, & \neg p, & \neg q\} \end{array}$$

Like in the previous example, we examine the minimal models on the basis of the statement of default applicability. Since there is no counter to Δ *in* M_2*, we do not need to consider this case. In* M_1*,* Δ' *(*$=\{not\_q\}$*) is not applicable because* Δ' *is incoherent (*$T \models_{M_1} not\_q \wedge q$*). In this example, the situation is different from the previous example, i.e.* $T \not\models_{MIN} q$.

This example provides us with another suggestion of the condition that Δ' is not applicable after Δ is applied. The next definition is based on this examination.

Definition 3.6 *Weakly Admissible hypotheses supported by minimal models A coherent hypotheses set* Δ *for a program* P *(*$T = P \cup \Delta$*) is weakly admissible iff* T *satisfies the condition:*

$\forall N \in MIN(T), \forall A \in HB,$
$T \models not\_A$ *and* $T \models_N A \Rightarrow \exists not\_A' \notin \Delta, T \models_N not\_A' \wedge A'.$

Weakly Admissible(P) denotes the class of all weakly admissible hypotheses set for a program P.

The next example is the last to be considered in formalizing the statement of default applicability.

Example 3.3 *Let us consider the following program and a hypothesis* $\Delta = \{not\_p\}$ *(Δ is coherent) which we would like to apply.*

$$p \leftarrow not\_q,$$

$$q \leftarrow not\_q, not\_r.$$

Here we show the minimal models for $T = P \cup \Delta$:

$$\begin{array}{llllll} M_1: & \{not\_p, & not\_q, not\_r, & p, & q, & \neg r\} \\ M_2: & \{not\_p, & \neg not\_q, not\_r, & \neg p, & \neg q, & \neg r\} \\ M_3: & \{not\_p, & not\_q, \neg not\_r, & p, & \neg q, & \neg r\} \\ M_4: & \{not\_p, & \neg not\_q, \neg not\_r, & \neg p, & \neg q, & \neg r\} \end{array}$$

Since there is no counter to Δ in M_2 and M_4, we do not need to consider these cases. In M_1, Δ' (=$\{not\_q, not\_r\}$) is not applicable because $T \models_{M_1} not\_q \wedge q$ (this is the same reason as in the previous example). However, there is no incoherency on Δ' (=$\{not\_q\}$) in M_3, which implies a counter to Δ.

If we can say that Δ' (=$\{not\_q\}$) is not applicable, Δ is justified in being applicable by the statement in previous section. Here, we consider models after Δ' is applied in addition to Δ, i.e. M_1 and M_3. We find that M_1 implies a counter, q, to Δ'. Furthermore, in M_1, Δ'' (=$\{not\_r\}$) is applicable because Δ'' is coherent. This means that we should recognize the counter to Δ' as the unavoidable one, since the counter to Δ' is derived by applying the applicable default Δ''. In other words, we conclude that in M_3 Δ' is not applicable after Δ is applied.

This investigation provides us with another suggestion of the conditions where Δ' is not applicable after Δ is applied.

Definition 3.7 *Acceptable hypotheses supported by minimal models*
A coherent hypotheses set Δ for a program P ($T = P \cup \Delta$) is acceptable iff T satisfies the condition:

$\forall N \in MIN(T), \forall A \in HB,$
if $T \models not\_A$ *and* $T \models_N A,$
then $\exists not\_A' \notin \Delta, T \models_N not\_A' \wedge A',$
otherwise $\exists NN (\neq N)$ *s.t.* $T \models_N not\_A' \Rightarrow T \models_{NN} not\_A',$
$\exists not\_A' \notin \Delta, T \models_N not\_A' \wedge T \models_{NN} A'$
and $\exists not\_A'$ *s.t.* $T \not\models_N not\_A', T \models_{NN} not\_A' \wedge \neg A'$

Acceptable(P) denotes the classes of all acceptable hypotheses set for a program P.

We show the inclusion of three class of conditions defined in this section.

Proposition 3.2 *For a program P, Admissible(P) $\subseteq$ Weakly Admissible(P) $\subseteq$ Acceptable(P).*

4 Expandability of Interpretations

In this section, we redefine the three conditions provided in the previous section in terms of the negatively supported interpretations and expandabilities first shown in [Brogi91].

First, we restate the basic concepts in [Brogi91]

Definition 4.1 *(Supported interpretations) Let P be a program. An interpretation M of P (i.e. $M \subseteq HB \cup not\_HB$) is a supported interpretation of P iff $H \subseteq M$ such that $M = T_{P \cup H} \uparrow \omega$.*

Given an interpretation I, I^+ stands for $I \cap HB$ and I^- stands for $I \cap not\_HB$. The supported interpretation $M(\Delta)$ such that $M(\Delta) = T_{P \cup \Delta} \uparrow \omega$ is called a negatively supported interpretation. We may express a negatively supported interpretation $M(\Delta)$ simply as M ($\Delta = M^-$). For an interpretation M, M is a coherent interpretation if M^- is coherent to the program.

At this point, we can give the natural correspondence between minimal models and negatively supported interpretations.

Proposition 4.1
1. *Let M be a negatively supported interpretation, $T = P \cup M^-$. For any negatively supported interpretation N s.t. $N^- \supseteq M^-$, there exists a minimal model $\mathcal{N}$ s.t. $\{not\_A | T \models_{\mathcal{N}} not\_A\} = N^-$, $\{A | T \models_{\mathcal{N}} A\} = N^+$. Especially, in the case of $N = M$, $\{not\_A | T \models_{\mathcal{N}} not\_A\} = \{not\_A | T \models not\_A\} = \{not\_A | T \models_{MIN} not\_A\} (= M^-)$, and $\{A | T \models_{\mathcal{N}} A\} = \{A | T \models A\} = \{A | T \models_{MIN} A\} (= M^+)$*

2. *For a minimal model $\mathcal{N}$ for $T = P \cup \Delta$, there is a negatively supported interpretation N such that $N^- = \{not\_A | T \models_{\mathcal{N}} not\_A\}$, $N^+ = \{A | T \models_{\mathcal{N}} A\}$. Especially, in the case of $\{not\_A | T \models_{\mathcal{N}} not\_A\} = \Delta$, $N^- = \{not\_A | T \models_{\mathcal{N}} not\_A\} = \{not\_A | T \models not\_A\} = \{not\_A | T \models_{MIN} not\_A\} (= \Delta)$, $N^+ = \{A | T \models_{\mathcal{N}} A\} = \{A | T \models A\} = \{A | T \models_{MIN} A\}$.*

The following series of definitions and theorems show rewriting of the conditions in the previous section in terms of negatively supported interpretations.

Definition 4.2 *[Brogi91] Admissible supported interpretation and models*
Let M be a negatively supported interpretation of a program P. M is an admissible supported interpretation *iff:*

$$\forall \text{ negatively supported interpretation } N\text{: } N^- \supseteq M^-,$$
$$M^- \cup N^+ : \text{incoherent} \Rightarrow M^+ \cup N^- : \text{incoherent}.$$

M is an admissible supported model if M is coherent. $ASM(P)$ denotes the class of all admissible supported models for P.

Theorem 4.1 *admissible hypotheses = admissible supported model*

1. *M is an admissible supported model $\Rightarrow M^-$ is admissible.*
2. *Δ is admissible $\Rightarrow$ a negatively supported interpretation $M(\Delta)$ is an admissible supported model.*

Proof: See the appendix.

Definition 4.3 *Acceptable Interpretations and Models*
Let M be a negatively supported interpretation of a program P. M is an acceptable interpretation *iff:*

> *$\forall$ negatively supported interpretation N: $N^- \supseteq M^-$, if $M^- \cup N^+$: incoherent then $\exists$ negatively supported interpretation NN: $NN^- \supseteq N^-$, $N^- - M^- \cup NN^+$: incoherent and $NN^- - N^- \cup NN^+$: coherent.*

M is an acceptable model if M is coherent. $AM(P)$ denotes the class of all acceptable models for P.

Theorem 4.2 *acceptable hypotheses = acceptable model*

1. *M is an acceptable model $\Rightarrow M^-$ is acceptable.*
2. *Δ is acceptable $\Rightarrow$ a negatively supported interpretation $M(\Delta)$ is an acceptable model.*

Proof: Same as the proof of theorem 4.1.

Definition 4.4 *weakly admissible supported interpretation and model*
Let M be a negatively supported interpretation of a program P. M is a weakly admissible supported interpretation *iff:*

$$\forall \textit{ negatively supported interpretation } N\text{: } N^- \supseteq M^-,$$
$$M^- \cup N^+ : \textit{incoherent} \Rightarrow N^- - M^- \cup N^+ : \textit{incoherent}.$$

M is an weakly admissible supported model if M is coherent. WASM(P) denotes the class of all weakly admissible supported models for P.

Theorem 4.3 *weakly admissible hypotheses = weakly admissible supported model*

1. *M is a weakly admissible supported model* $\Rightarrow M^-$ *is weakly admissible.*
2. Δ *is weakly admissible* $\Rightarrow$ *a negatively supported interpretation* $M(\Delta)$ *is a weakly admissible supported model.*

Proof: Same as the proof of theorem 4.1.

Before closing this section, we define special cases of acceptable interpretations which are used in the next section.

Definition 4.5 *Strictly Acceptable Interpretations and Models*
Let M be a negatively supported interpretation of a program P. M is a strictly acceptable interpretation *iff:*

> $\forall$ *negatively supported interpretation N:* $N^- \supseteq M^-$,
> *if* $M^- \cup N^+$: *incoherent then* $\exists$ *negatively supported interpretation NN:* $NN^- \supseteq N^-$, $N^- - M^- \cup NN^+$: *incoherent and* $\forall NNN^- \supseteq NN^-$, $NN^- - N^- \cup NNN^+$: *coherent.*

M is a strictly acceptable model if M is coherent. SAM(P) denotes the class of all strictly acceptable models for P.

Corollary 4.1 *For a program P,* $ASM(P) \subseteq WASM(P) \subsetneq SAM(P) \subseteq AM(P)$

5 Acceptability from Operational Perspectives

In this section, we show that the notions of acceptability in negatively supported interpretations in the previous section are equivalent to operational semantics based on the principle of negation as abductive hypothesis defined by Dung [Dung91], Kakas and Mancarella [Kakas91], Kakas, Kowalski and Toni [Kakas92]. Therefore, we conclude that the class of acceptability supported by minimal models is equivalent to the class of operational semantics.

Here, we present the ideas of operational semantics based on the principle of negation as abductive hypotheses. Eshghi [Eshghi89] firstly showed the treatment of negation as hypothesis to reformalize the stable model semantics by Gelfond and Lifschitz [Gelfond88]. The main idea of this approach is to give the semantics of a program P through the notion of derivability in the extended program $P \cup \Delta$ (Δ is a set of negative (or abductive) propositions of the form not_A).

In the sequel, we recall the definitions in [Kakas92] of operational semantics according to our terminologies. Given a set Δ ($\subseteq$ not_HB), a set $\mathcal{A}$ ($\subseteq$ not_HB) *attacks* Δ if $P \cup \mathcal{A} \vdash A$ for some $not\_A \in \Delta$.

In [Eshghi89], a set of acceptable hypotheses Δ is coherent and Δ should result in a total model, i.e. $P \cup \Delta$ derives A or not_A for each proposition A. Dung [Dung91] relaxed the condition of totality to the maximality of Δ. But he needed another condition on acceptability in his *preferential semantics*, that is the *admissibility* of hypotheses Δ:

$$\forall \mathcal{A}\text{: } \mathcal{A} \text{ attacks } \Delta,\ \Delta \text{ attacks } \mathcal{A}.$$

Then, to save unintuitive examples in the preferential semantics, Kakas and Mancarella [Kakas91] consider one other acceptability condition on hypotheses in their *stable theory*, that is the *weakly stability* of hypotheses Δ:

$$\forall \mathcal{A}\text{: } \mathcal{A} \text{ attacks } \Delta,\ \mathcal{A} \cup \Delta \text{ attacks } \mathcal{A} - \Delta.$$

In the next theorem, we restate the result in [Brogi91].

Theorem 5.1 *[Brogi91] admissible supported model = admissible*

1. *a coherent hypotheses set Δ is admissible $\Rightarrow$ a negatively supported interpretation $M(\Delta)$ is an admissible supported model.*

2. M is an admissible supported model $\Rightarrow M^-$ is admissible.

In addition to the above theorem, we can state the next theorem, which gives the equivalence between weakly admissible supported models and weakly stable hypotheses of Kakas and Mancarella.

Theorem 5.2 *stable theory = weakly admissible supported model*

1. *a coherent hypotheses set Δ is weakly stable $\Rightarrow$ a negatively supported interpretation $M(\Delta)$ is a weakly admissible supported model.*

2. *M is a weakly admissible supported model $\Rightarrow M^-$ is weakly stable.*

Proof: See the appendix.

In [Kakas91], to save unintuitive examples in stable theories, they provide an alternative condition of acceptability in their *acceptability semantics*, that is the *acceptability* of hypotheses set Δ: for some initial hypotheses set Δ_0,

Δ is acceptable to Δ_0 iff $\forall\mathcal{A}$: $\mathcal{A}$ attacks Δ, $\mathcal{A}$ is not acceptable to $\Delta \cup \Delta_0$.

In the sequel, we only consider whether Δ is acceptable to $\emptyset$. As shown in [Toni92], we can unfold the recursion using the above definition of acceptability once:

Δ is acceptable to $\emptyset$
iff $\forall\mathcal{A}$: $\mathcal{A}$ attacks Δ, $\exists\mathcal{D}$: $\mathcal{D}$ attacks $\mathcal{A} - \Delta$ s.t. $\mathcal{D}$ is acceptable to $\mathcal{A} \cup \Delta$.

Finally, we can state the next theorem which gives the equivalence between acceptable models and acceptable hypotheses to $\emptyset$.

Theorem 5.3 *acceptability semantics = acceptable model*

1. *Δ is acceptable to $\emptyset \Rightarrow$ a negatively supported interpretation $M(\Delta)$ is an acceptable model.*

2. *M is a strictly acceptable model $\rightarrow M^-$ is acceptable to $\emptyset$.*

Proof: See the appendix.
The following example shows the case where M^- is not acceptable to $\emptyset$ though M is not a strictly acceptable model but acceptable model.

Example 5.1 *Let us consider the following program and a hypothesis* $\Delta = \{not\_p(0)\}$ *(Δ is coherent).*

$$p(X) \leftarrow not\_p(s(X)),$$

$$q(0) \leftarrow .$$

Hypothesis Δ is an acceptable model, but not a strictly acceptable model. Moreover, Δ is not acceptable to $\emptyset$ because we need infinite recursive calls of acceptability definition.

6 Concluding Remarks

We consider an intuitive criterion to classify acceptable hypotheses as proper negation by default, and examine the small programs on the basis of this criterion. Then, we provide three variants of conditions for acceptable negative hypotheses in terms of minimal models. The conditions are rewritten in negatively supported interpretations. We have proven that the class of operational semantics is equivalent to the class of conditions. The table shows these equivalences among conditions:

Framework	Minimal Models	Expandability	Operational
	Acceptable	Acceptable	Acceptable
	Weakly Admissible	Weakly Admissible	Weakly Stable
	Admissible	Admissible	Admissible

Now, we have two directions in which to proceed with future works. The first direction is concerned with the way in which the minimal models are discussed. In Section 3, we discuss minimal models by regarding not_A as a positive hypothesis. If we may treat not_A as $\neg LA$ in autoepsitemic logic (L is a modal operator, and LA means that "A is believed"), we might provide the conditions in autoepistemic logic terminologies. Another topic is the maximality of hypotheses, which is not mentioned in this paper. Since maximality is an important measure, as stated in [Dung91] and [Kakas91], we would like to incorporate maximality into our formalization.

The second direction is related to extensions of languages. We would like to specify the semantics of extended logic programs and abductive logic programs which have positive and negative hypotheses (with or without integrity constraints) in the same way as in this work.

Acknowledgment

I am grateful to Ken Satoh for the suggestion of relation to autoepistemic logic, and to Robert A. Kowalski for the indication of significance of maximality.

References

[Brogi91] Brogi, A., Mancarella, P., Lamma, P., Mello, P., Normal Logic Programs as Open Logic Programs, to appear in *Proc. of JISCLP'92*.

[Dung91] Dung, P. M., Negation as Hypotheses: An Abductive Foundation of Logic Programming, *Proc. of ICLP'91*, pp. 3 – 17 (1991).

[Eshghi89] Eshghi, K., Kowalski, R. A., Abduction Compared with Negation by Failure, *Proc. of ICLP'89*, pp. 234 – 254 (1989).

[Gelfond88] Gelfond, M., Lifschitz, V., The Stable Model Semantics for Logic Programming, *Proc. of ICLP'88*, pp. 1070 – 1080, 1988.

[Kakas90] Kakas, A. C., Mancarella, P., Generalized Stable Models: A Semantics for Abduction, *Proc. of ECAI'90*, pp. 385 – 391 (1990).

[Kakas91] Kakas, A. C., Mancarella, P., Stable Theories for Logic Programs, *Proc. of ILPS'91*, pp. 85 – 100 (1991).

[Kakas92] Kakas, A. C., Kowalski, R. A., Toni, F., Abductive Logic Programming, Imperial College Technical report, (1992).

[Przymusinski91] Przymusinski, C. T., Semantics of Disjunctive Logic Programs and Deductive Databases, *Proc. of DOOD'91*, LNCS566, pp. 85 – 107 (1991).

[Toni92] Toni, F., Kowalski, R. A., A Proof Procedure for Acceptability Semantics, private communication, (1992)

Appendix

Proof of theorem 4.1

By proposition 4.1, we show the following correspondences between minimal models and negatively supported interpretations: for $not\_A \in \Delta$,

$$T \models not\_A \Leftrightarrow not\_A \in M^-,$$

$$T \models_N A \Leftrightarrow A \in N^+,$$

and

$$not\_A' \notin \Delta, T \models_N not\_A' \Leftrightarrow not\_A' \in N^- - M^-,$$

$$T \models_{MIN} A' \Leftrightarrow A' \in M^+.$$

So the condition

> $\forall N \in MIN(T), \forall A \in HB,$
> $T \models not\_A$ and $T \not\models_N A \Rightarrow \exists not\_A' \notin \Delta, T \models_N not\_A'$ and $T \models_{MIN} A'$

is equivalent to the condition

> $\forall N : N^- \supseteq M^-,$
> $M^- \cup N^+$: incoherent $\Rightarrow M^+ \cup N^- - M^-$: incoherent.

Since M is coherent, $M^+ \cup N^- - M^-$ is incoherent iff $M^+ \cup N^-$ is incoherent. □

Proof of theorem 5.2

1) Assume $M(\Delta)$ is not a weakly admissible supported model. Then $\exists N : N^- \supseteq M^-$ s.t. $M^- \cup N^+$ is incoherent and $N^- - M^- \cup N^+$ is coherent, that is $\forall not\_A \in N^- - M^-$ s.t. $P \cup N^- \not\vdash A$. This contradicts the weakly stability of Δ because we can regard that $N^- = \mathcal{A} \cup \Delta$.
2) It is enough to consider the case where $N^- = \mathcal{A}$ s.t. $\mathcal{A} \supseteq \Delta$. □

Proof of theorem 5.3

1) If Δ is incoherent, then $\exists \mathcal{A}$, $\mathcal{A}$ attacks Δ and $\mathcal{A} \subseteq \Delta$. So, there is no $\mathcal{D}$ s.t. $\mathcal{D}$ attacks $\mathcal{A} - \Delta$ because $\mathcal{A} - \Delta = \emptyset$. This contradicts the acceptability of Δ.

Assume that $M(\Delta)$ is not an acceptable model, i.e. $\exists N : N^- \supset M^-$ s.t. $M^- \cup N^+$ is incoherent, $\forall NN : NN^- \supseteq N^-$, $N^- - M^- \cup NN^+$: coherent

or $NN^- - N^- \cup NN^+$: incoherent. By regarding $N^- = \mathcal{A}$ and $NN^- = \mathcal{D}$ we obtain that NN^- is acceptable to N^- because Δ is acceptable to $\emptyset$ if $N^- - M^- \cup NN^+$ is incoherent. However, this is a contradiction since NN^- attacks $NN^- - N^-$ but there is no $\mathcal{D}$ which attacks $NN^- - (NN^- \cup N^-) = \emptyset$.
2) Assume that M^- is not acceptable to $\emptyset$, i.e. $\exists \mathcal{A}$, $\mathcal{A}$ attacks M^-, $\forall \mathcal{D}$, $\mathcal{D}$ is not acceptable to $\mathcal{A} \cup M^-$. Let us consider the case where $\mathcal{A} \supset M^-$; otherwise let us consider $\mathcal{A} \cup M^-$ as a new $\mathcal{A}$. Since M is acceptable model, $\exists NN^-$ attacks $\mathcal{A} - M^-$. So, NN^- is not acceptable to $\mathcal{A}$. Therefore, $\exists \mathcal{A}'$ attacks $NN^- - \mathcal{A}$, $\mathcal{A}'$ is acceptable to NN^-. However, this contradicts the fact that $\forall NNN^- \supseteq NN^-$, $NN^- - N^- \cup NNN^+$: coherent (M is a strictly acceptable model), since $\mathcal{A}' \cup NN^-$ attacks $NN^- - \mathcal{A}$. $\Box$