Partial Models and Three-Valued Models in Logic Programs with Negation

Domenico Saccà†
Dipartimento di Sistemi
Università della Calabria
Rende, Italy

Carlo Zaniolo
MCC
Austin, Texas, USA
carlo@mcc.com

Abstract

Much of the current work in non-monotonic logic pursues the generalization of concepts such as well-founded models and stable models using three-valued logic. This approach is also effective in dealing with incomplete and undefined information that is frequently found in knowledge bases. However, it also suffers from drawbacks, including the fact that, in multi-valued logic, there is more than one meaningful way to assign a meaning to rules in a program. In this paper, we present a reconstruction of theory of negation in logic rules which deals with incompleteness and undefinedness using the standard two-valued logic. Simple extensions of the notion of unfounded sets are used to define the concept of partial models and the notions of partial well-founded models and partial stable models. We prove that the partial stable models so defined are equivalent to the three-valued stable models proposed by Przymusinski. On semantic grounds, however, these models suffer from serious drawbacks caused by with their inability to enforce the principle of minimal undefinedness and thus we argue for the need of a stricter semantics.

1. Introduction

The problem of providing a formal semantics to programs where rules contain negated goals represents a key of research issue in areas such as logic programming, non-monotonic reasoning and deductive databases. Significant progress has been made recently on this topic, largely as a

† This author's work was done in part under contract by MCC and was also supported by the Italian National Research Council as part of the project "Sistemi Informatici e Calcolo Parallelo".

result of a renewed interest in deductive databases. For instance, the concept of *stratified programs* that was introduced only a few years ago [ABW, CH, N, V1] is now regarded as a standard notion, efficiently supported in systems such as NAIL! [U] and LDL [Ceta], as part of their "bottom-up" execution strategy. However, to remove the limitations of stratified programs, the more general notions of *well-founded models* [VRS, V2] and *stable models* [GL] were proposed.

While these concepts were initially defined for the domain of total interpretations, current research is focusing on partial models, where facts need not be classified as either true or false, but, rather, can be left undefined. Thus, partial models are capable of dealing with incomplete or locally inconsistent information and provide a very powerful and flexible device for generalizing the formal semantics of logic programs with negation. Thus, generalized concepts of stable models and well-founded are proposed in [P3], such that these kinds of models exist for all programs. These approaches, however, rely on three-valued logic for their formal developments, and, therefore, have been the object of criticisms [YoYu], which, e.g., point out that there there is no unique way to generalize the notion of logical implication in multi-valued logic.

In this paper, we present a reconstruction of the theory of negation in logic rules that deals with incompleteness and undefinedness using the standard two-valued logic. Simple extensions of the notion of unfounded sets are used to introduce the concept of partial model, and then, to refine this notion into well-founded models and stable models.

2. Preliminary Definitions

Let us start by defining our language (Horn clauses plus negated goals in rules) and basic concepts and notation [L,U].

A *term* is a variable, a constant, or a complex term of the form $f(t_1, \ldots, t_n)$, where $t_1, \ldots, t_n$ are terms. An *atom* is a formula of the language that is of the form $p(\mathbf{t})$ where p is a predicate symbol of a finite arity (say n) and $\mathbf{t}$ is a sequence of terms of length n (*arguments*). A *literal* is either an atom (*positive literal*) or its negation (*negative literal*). An atom A, and its negation, i.e., the literal $\neg A$, are said to be the *complement* of each other. In general, if B is a literal, then $\neg B$ denotes the complement of B.

A *rule* r is a formula of the language of the form

$$Q \leftarrow Q_1, \ldots, Q_m .$$

where Q is a atom (*head* of the rule) and $Q_1, \ldots, Q_m$ are literals (*goals* of the rule). Let $H(r)$ and $B(r)$ represent, respectively, the head of r and the set of all goals of r. Whenever no confusion arises, we shall also see $B(r)$ as a conjunction of goals.

A term, atom, literal or rule is *ground* if it is free of variables. A ground rule with no goals is a *fact*. A *logic program* is a set of rules. A rule without negative goals is called positive (a Horn clause); a program is called *positive* when all its rules are positive.

Given a logic program P, the *Herbrand universe* for P, denoted H_P, is the set of all possible ground terms recursively constructed by using constants and function symbols occurring in P. The *Herbrand Base* of P, denoted B_P, is the set of all possible ground atoms whose predicate symbols occur in P and whose arguments are elements of the H_P. A *ground instance* of a rule r in P is a rule obtained from r by replacing every variable X in r by $\phi(X)$, where ϕ is a mapping from all variables occurring in r to terms in the Herbrand universe. The set of all ground instances of r are denoted by *ground* (r); accordingly, *ground* (P) denotes $\cup_{r \in P}$ *ground* (r).

Let X be a set of ground literals; then $\neg X$ denotes the set $\{\neg A \mid A \in X\}$, X^+ (resp., X^-) denotes the set of all positive (resp., negative) literals in X. Finally, $\overline{X}$ denotes all elements of the Herbrand Base which do not occur in X, i.e., $\overline{X} = \{A \mid A \in B_P$ and neither A nor $\neg A$ is in $X\}$.

Definition 1. Interpretation: Let P be a logic program. Given a subset I of $B_P \cup \neg B_P$, I is a *(partial) interpretation* of P if it is *consistent*, i.e., it does not contain two elements which are the complement of each other. Moreover, if $I^+ \cup \neg I^- = B_P$, the interpretation I is called *total.* □

A ground rule r is made *true* by a total interpretation I, if $H(r)$ is in I^+ or if $B(r)$ is not contained in I.

Definition 2. Total Model: Let P be a logic program. A *total model*, M of P is a total interpretation of P that makes each ground instance of each rule in P true. □

Definition 3. Minimal Model: A *minimal model* M of P is a total model for which there exists no other total model N such that N^+ is a proper subset of M^+. □

It is well-known that a positive program has a unique minimal model which represents its natural meaning. In this case, the set of positive literals in the minimal model can be determined using a least fixpoint computation. This computation is based on the *immediate consequence transformation* $T_P : 2^{B_P \cup \neg B_P} \rightarrow 2^{B_P \cup \neg B_P}$, where for each X in $2^{B_P \cup \neg B_P}$, $T_P(X) = \{A \mid A = H(r), r \in ground(P)$ and $B(r) \subseteq X\}$. The transformation T_P is monotone and continuous in the complete lattice $< 2^{B_P \cup \neg B_P}, \subseteq >$ and, then, the least fixpoint of T_P, denoted by $T_P^\infty(\emptyset)$, exists and coincides with $\cup_{i \geq 0} T_P^i(\emptyset)$, where $T_P(\emptyset) = \emptyset$ and $T_P^i(\emptyset) = T_P(T_P(\emptyset))$ [T, L]. Let $M_T^+ = T_P^\infty(\emptyset)$. If P is a positive program then $M_T = M_T^+ \cup \neg(B_P - M_T^+)$ is the minimal model of P. (Note that if P is not positive then M_T is an interpretation but not necessarily a model.)

In case of non-positive programs, the existence of a unique minimal model is not guaranteed. Therefore, the issue of what model to base the semantics of the program becomes much more complex. The stable model semantics, introduced in [GL] and defined next, represents an interesting solution to the problem which is based on a particular usage of T_P.

Given a program P and a total model M for P, the *positive version* of P w.r.t. M, denoted P_M, is the positive program obtained from *ground*(P) by deleting (a) each rule that has a negative goal $\neg A$ such that $\neg A$ is not in M^-, and (b) all negative goals from the remaining rules.

Definition 4. Stable Total Model: Let M a total model for a program P. Then if $T_{P_M}^{\infty}(\emptyset) = M^+$, M is said to be a *stable model* for P. □

The condition $T_{P_M}^{\infty}(\emptyset) = M^+$, is called the *stability condition* for P w.r.t. M. It is easy to see that if M_T is a model then M_T is both a stable model and the unique minimal model, thus stable model semantics is an immediate extension of positive program semantics. The following important property is proven in [GL]:

FACT 1. *Stable models are minimal models* □

We point out that a stable model is not necessarily the unique minimal model of a program or even the unique stable model.

Example 1. Consider the following program having 0-arity predicate symbols:

$u \leftarrow \neg v.$

$v \leftarrow \neg u.$

There are three total models: $M_1 = \{u, \neg v\}$, $M_2 = \{v, \neg u\}$, $M_3 = \{u, v\}$. Only M_1 and M_2 are stable; furthermore, they are the only minimal models. □

The issue of multiple stable models was discussed in [SZ1] and [GPZ] where it was shown that this is a manifestation of the non-determinism that is implicit in the stable model semantics.

3. Partial Models, Stable Models and Well-Founded Models

We begin with some useful definitions:

Definition 5: Blocked and Inapplicable rules: Given $X \subseteq 2^{B_P \cup \neg B_P}$ and a rule r in *ground*(P), r is

(a) *blocked* w.r.t. X if there exists an element A in $B(r)$ such that $\neg A$ is in X;

(b) *inapplicable* w.r.t. X if $B(r) \subseteq X$ does not hold. □

Thus, while a blocked rule contains a goal with is false (w.r.t. X) a inapplicable rule contain some goal that is not true. In the case that X is an interpretation, the predicates which do not occur in X (i.e., those in $\overline{X}$) are not known to be true or false and, then, they can be thought of as "undefined facts"; therefore, a non-blocked inapplicable rule can become applicable after some assignment of values to undefined facts.

Let us now define the important notion of unfounded set [VRS], and the related notion of assumption set.

Definition 6. Unfounded Sets and Assumption Sets. Let $X \subseteq 2^{B_P \cup \neg B_P}$ and $Y \subseteq B_P$.

(a) Y is an *unfounded set w.r.t.* X if either (i) Y is empty or (ii) for each A in Y, every rule r with head A in *ground*(P) is either blocked w.r.t. X or $B(r) \cap Y \neq \varnothing$.

(b) Y is an *assumption set w.r.t.* X if either (i) Y is empty or (ii) for each A in Y, every rule r with head A in *ground*(P) is either inapplicable w.r.t. X or $B(r) \cap Y \neq \varnothing$. □

Obviously, if X is an interpretation, every unfounded set is an assumption set but the converse is not true. Moreover, for total interpretations, the two concepts coincide.

Example 2. Consider the program of Example 1:

$u \leftarrow \neg v.$

$v \leftarrow \neg u.$

and its interpretation $M_1 = \{u, \neg v\}$. We have that $\{u\}$ is not an assumption set w.r.t. M_1 since the goal of the first rule is in $M_1 - \{u\}$. Note that $\{u, v\}$ is an assumption set (but not an unfounded set) w.r.t. the empty set. Finally $\{\neg v\}$ is an unfounded set w.r.t. M_1. □

Let us now extend the definition of model to the domain of partial interpretations.

Definition 7. Partial Models and Total Models: An interpretation M of a program P is a *model* of P if M^- is an unfounded set w.r.t. M. Moreover, if M is a total interpretation then M is called a *total model.* □

We next show that Definition 7 is consistent with Definition 4 and that partial models have the the following intuitive properties:

PROPOSITION 1. *The following properties hold:*

(a) *An interpretation M of P is a model of P if and only if for each $\neg A$ in M^-, every rule r in ground(P) with $H(r) = A$ is blocked w.r.t. M;*

(b) *for total models Definitions 4 and 7 are equivalent;*

(c) *every model of a program P is a subset of some total model.*

PROOF. (*a*) Suppose first that M be an interpretation of P such that for each $\neg A$ in M^-, every rule r in *ground*(P) with $H(r) = A$ is blocked w.r.t. M. Then, by Definition 6(a), M^- is an unfounded set w.r.t. M. Suppose now that M is a model of P, thus M^- is an unfounded set w.r.t. M. Then, by Definition 6(a), for each $\neg A$ in M^-, every rule r in *ground*(P) with $H(r) = A$ is either (1) blocked w.r.t. M or (2) $B(r) \cap \neg M^- \neq \varnothing$. It is easy to see that also in the case (2) r is blocked w.r.t. M.

(*b*) Suppose that M is a total interpretation of P. Hence, for each rule r in *ground*(P), if $H(r)$ not in M^+ then $\neg H(r)$ is in M^-. Therefore, Part (*a*) of this proposition can be rephrased as follows: M is a total model of P (according to Definition 7) if and only if for each rule r in *ground*(P), either $H(r)$ is in M^+ or r is blocked w.r.t. M. But, by Definition 5(b), if r is blocked w.r.t. M then $B(r)$ is not contained in M; so, in both cases, r is made true by M, thus Definitions 4 and 7 are equivalent.

(c) Let M be a model of P. Consider the set $N = M \cup \overline{M}$. In order to prove the proposition it is sufficient to show that N is a total model. By construction, N is consistent and every element of B_P occurs in N; therefore N is a total interpretation of P. By Definition 7, M^- is an unfounded set w.r.t. M; therefore, as $M \subseteq N$, M^- is an unfounded set also w.r.t. N. But $N^- = M^-$; so, by Definition 7, N is a total model of P. □

Note that the previous definition of partial model M is stricter than that used in [VRS], which would accept any subset of some total model.

Example 3. Consider the following program:

$p(a) \leftarrow \neg p(b).$

$p(c) \leftarrow p(b).$

$p(d) \leftarrow p(c).$

$\{p(a), \neg p(b), \neg p(c)\}$ is a partial model and is a subset of the total minimal model $\{p(a), \neg p(b), \neg p(c), \neg p(d)\}$. Note that $\{p(a), \neg p(c)\}$ is also a subset of that above total model but is not a partial model since, by assigning the value "true" to $p(b)$, we get a contradiction with $\neg p(c)$. □

As a logic program may have several models, it is crucial to introduce some criteria to recognize desirable models — i.e., models which are closer to the "intended" meaning of the program. As discussed in [YoYu], the three key properties considered highly desirable by researchers in this area are

(1) *consistency,*

(2) *justifiability,*

(3) *minimal undefinedness.*

The first principle is guaranteed by our definition of partial model. The second principle is found under several names in the work of several authors (e.g., the notion justifiability used in [YoYu] is similar to that of 'genuine supportedness' [Do] or that of 'foundedness' [SZ]). Basically it prescribes that every positive conclusion be demonstratable by reasoning that follows the orientation of the rules. The third principle says that the number of undefined facts should be reduced as much as possible. The definition of stable models, in the domain of total interpretations, can be viewed as a direct implementation of these principles.

We can now introduce the concept of partial stable models (or P-stable model for short), according to the principles above and following observations:

(1) adding the complements of facts from an unfounded set to a model, reduces undefinedness while preserving consistency and justifiability,

(2) adding facts from an assumption set to a model might violate the justifiability principle; adding their complements might compromise consistency.

Definition 8. P-Stable Models. Let P be a program and M a model of it. Then M is *P-stable* if

(a) no non-empty subset of M^+ is an assumption set w.r.t. M;

(b) every non-empty subset of $\overline{M}$ is an assumption set w.r.t. M, while it is not an unfounded set w.r.t. M. □

Example 4. Consider now the following program:

$a \leftarrow \neg a.$

$b \leftarrow \neg c, d.$

$c \leftarrow d.$

$d \leftarrow b.$

We have that the $M = \{\neg b, \neg c, \neg d\}$ is the unique P-stable model. Note that $\overline{M} = \{a\}$ is an assumption set w.r.t. to M but not an unfounded set, as neither the first rule is blocked nor a is one of its goals.

For the program of Example 1, we have three P-stable models: one is the empty set, the others are $\{u, \neg v\}$ and $\{\neg u, v\}$. □

As far as the principle of justifiability is concerned, the next result shows that the stability condition as for total stable models holds for P-stable models as well.

LEMMA 1. *Let P be a program and M be a model of it. If the stability condition, $T_{P_M}^{\infty}(\emptyset) = M^+$, holds then (a) every non-empty subset of M^+ is an assumption set w.r.t. M and (b) every subset of $\overline{M}$ is an assumption set w.r.t. M. Viceversa, if (a) and (b) hold, so does the stability condition.*

PROOF. Let M be a model and P_M be the positive version of P We first prove the following technical result:

CLAIM 1. *If every subset of $\overline{M}$ is assumption set w.r.t. M then for each A in $T_{P_M}^{\infty}(\emptyset)$, A is in M and there exists a rule r in ground(P) such that $H(r) = A$ and $B(r) \subseteq M$.*

PROOF. Let $T^0 = \emptyset$ and $T^k = T_{P_M}(T^{k-1})$, $k > 0$. By definition of $T_{P_M}^{\infty}(\emptyset)$, A is in $T_{P_M}^{\infty}(\emptyset)$ if and only if A is in T^k, for some $k > 0$. Therefore, in order to prove the claim, it is sufficient to show that for each k, $k \geq 0$, and for each A in T^k, A is in M and there exists a rule r in *ground*(P) such that $H(r) = A$ and $B(r) \subseteq M$. We proceed by induction on k. The claim trivially holds for $k = 0$ (basis of the induction). Let $k > 0$ and A be any element in T^k. By definition of immediate consequence transformation, there is a rule r in P_M such that $A = H(r)$ and

$B(r) \subseteq T^{k-1}$. By inductive hypothesis, $T^{k-1} \subseteq M^+$; so $B(r) \subseteq M^+$ as well. Let $\hat{r}$ be the rule in *ground*(P) from which r has been derived. By definition of positive version, $H(\hat{r}) = A$, $B(\hat{r})^+ = B(r)$ and $B(\hat{r})^- \subseteq M^-$; so $B(\hat{r}) \subseteq M$. Hence, $\hat{r}$ is not inapplicable w.r.t. M. We observe that: (1) $\neg A$ can not be in M^- by Part (a) of Proposition 1, and (2) A can not be in $\overline{M}$ because $\{A\}$ is not an assumption set w.r.t. M. It follows that A is in M^+, thus $T_{P_M}^{\infty}(\emptyset) = T^k \subseteq M^+$. But we have also shown that there exists a rule $\hat{r}$ in *ground*(P) such that $H(\hat{r}) = A$ and $B(\hat{r}) \subseteq M$. So also Part (b) of the claim is proved. □

(If part) Suppose that no non-empty subset of M^+ is an assumption set w.r.t. M and every subset of $\overline{M}$ is an assumption set w.r.t. M. By Part (a) of Claim 1, $T_{P_M}^{\infty}(\emptyset) \subseteq M^+$. Let $X = M^+ - T_{P_M}^{\infty}(\emptyset)$. We prove that X is empty by contradiction. Let A be an element in X. Since X is not an assumption set w.r.t. M by hypothesis, there exists at least one rule r in *ground*(P) with $H(r) = A$ such that both r is not inapplicable (i.e., $B(r) \subseteq M$) and $B(r) \cap X = \emptyset$. Hence, $B(r) \subseteq (M - X)$. As $B(r) \subseteq M$, by definition of positive version, the rule $\hat{r}$ obtained from r by removing negative goals is in P_M and $B(\hat{r}) \subseteq (M^+ - X) = T_{P_M}^{\infty}(\emptyset)$. Therefore, by definition of immediate consequence transformation, $A = H(r)$ is in $T_{P_M}^{\infty}(\emptyset)$ (a contradiction). It follows that X is empty and, then, $M^+ = T_{P_M}^{\infty}(\emptyset)$.

(Only-if part) Suppose now that $M^+ = T_{P_M}^{\infty}(\emptyset)$. Let X be any non-empty subset of $\overline{M}$ and A be any element in X. Let r be any rule in *ground*(P) with $H(r) = A$. We show by contradiction that r is inapplicable w.r.t. M. Suppose then that $B(r) \subseteq M$. By definition of positive version, the rule $\hat{r}$ in P_M corresponding to r is such that $B(\hat{r}) \subseteq M^+$. So A is in $T_{P_M}(M^+)$. But $M^+ = T_{P_M}(M^+)$ by definition of fixpoint; hence, by definition of T_{P_M} A is in M^+ (a contradiction). It follows that r is indeed inapplicable w.r.t. M and, then, X is an assumption set w.r.t. M. Hence, every non-empty subset of $\overline{M}$ is an assumption set w.r.t. M. Let us now prove by contradiction that every non-empty subset of M^+ is not an assumption set w.r.t. M. Suppose then that $X \subseteq M^+$ is an assumption set w.r.t. M. Let A be any element in X. Since A is in M^+, by Part (b) of Claim 1 there exists a rule r in *ground*(P) such that $h(r) = A$ and $B(r) \subseteq M$ - a contradiction with the assumption that X is an assumption set w.r.t. M. This concludes the proof. □

THEOREM 1. *Let P be a program and M be a P-stable model of it. Then the stability condition, $T_{P_M}^{\infty}(\emptyset) = M^+$, holds.*

PROOF. By definition of P-stable model, every non-empty subset of M^+ is not an assumption set w.r.t. M and every subset of $\overline{M}$ is an assumption set w.r.t. M. Hence, by Lemma 1, the stability condition holds for M. □

Considering Lemma 1 and Theorem 1, P-stable models can be also redefined as follows:

FACT 2. *Let P be a program and M be a model of it. Then M is P-stable if and only if both the stability condition holds for P w.r.t. M and no non-empty subset of $\overline{M}$ is an unfounded set w.r.t. M.* □

We next show that P-stable models have an interesting fixpoint characterization. To this end, we observe that, given a program P and $X \subseteq 2^{B_P \cup \neg B_P}$, the union of all unfounded sets w.r.t. X, denoted by $U_P(X)$, is also an unfounded set w.r.t. X. Let $W_P(X) = T_P(X) \cup \neg U_P(X)$. As both T_P and U_P are monotonic and continuous, W_P is a monotonic, continuous transformation in the complete lattice $< 2^{B_P \cup \neg B_P}, \subseteq>$ and, then, the least fixpoint of W_P exists, coincides with $\cup_{i \geq 0} W_P{}^i(\emptyset)$, where $W_P{}^0(\emptyset) = \emptyset$ and $W_P{}^i(\emptyset) = W_P(W_P{}^{i-1}(\emptyset))$, for $i > 0$, and is denoted by $W_P{}^\infty(\emptyset)$. Moreover, $W_P{}^\infty(\emptyset) = W_P(W_P{}^\infty(\emptyset))$. As it will be discussed later in this section, the least fixpoint is called the *well-founded model* [VRS] and is actually a P-stable model.

LEMMA 2. *Let P be a logic program and I be an interpretation of it. Then I is a fixpoint of W_P if and only if I is a model of P such that both $I^+ = T_P(I)$ and every non-empty subset of $\overline{I}$ is an assumption set w.r.t. I but not an unfounded set.*

PROOF. *(If part)* Let I be a model of P such that $I^+ = T_P(I)$ and no non-empty subset of $\overline{I}$ is an unfounded set w.r.t. I. By Definition 7, $I^- \subseteq U_P(I)$. Consider now any subset X of M^+. Let A be any element in X. Since $I^+ = T_P(I)$, there exists a rule r in *ground*(P) such that $B(r) \subseteq I$. Therefore, X is not an unfounded set w.r.t. I, thus $I^+ \cap U_P(I) = \emptyset$. Also $\overline{I} \cap U_P(I) = \emptyset$ by hypothesis. It follows that $I^- = U_P(I)$ and, then, $I = W_P(I)$.

(Only-if part) Let $I = W_P(I)$. Then $I^+ = T_P(I)$ and $I^- = U_P(I)$. From the latter equality, we derive that both I^- is an unfounded set w.r.t. I and $\overline{I} \cap U_P(I) = \emptyset$. It follows that I is a model of P by Definition 7 and no non-empty subset of $\overline{I}$ is an unfounded set w.r.t. I. Consider now any non-empty subset X of $\overline{M}$. Let A be any element in X. Since A is not in $T_P(I)$, every rule in *ground*(P) with $H(r) = A$ is either inapplicable w.r.t. M or one of the goals of r, say B, is in X. But also in the latter case r is inapplicable w.r.t. M for B is in $\overline{M}$. Therefore, X is not an assumption set w.r.t. I and this concludes the proof. □

THEOREM 2. *Every P-stable model of a logic program P is a fixpoint of W_P.*

PROOF. Let M be a P-stable model for P. Consider $W_P(M) = T_P(M) \cup \neg U_P(M)$. Since no non-empty subset of $\overline{M}$ is an unfounded set w.r.t. M, $M^- = U_P(M)$. We first prove that $M^+ = T_P(M)$ by contradiction. Suppose first that A is an element in $T_P(M) - M^+$. By definition of T_P, there exists a rule r in *ground*(P) such that $H(r) = A$ and $B(r) \subseteq M$; so $\{A\}$ is not an assumption set w.r.t. M — a contradiction with the hypothesis that M is P-stable (see Condition *b* of Definition 8). It follows that $T_P(M) \subseteq M^+$. Suppose now that A be an element in

$M^+ - T_P(M)$. By definition of T_P, every possible rule in *ground*(P) with $H(r) = A$ is inapplicable. Hence, $\{A\}$ is an assumption set w.r.t. M — again a contradiction with the hypothesis that M is P-stable (see Condition *a* of Definition 8). So we also have $M^+ \subseteq T_P(M)$ and, then, $M^+ = T_P(M)$. We now observe that every non-empty subset of $\overline{M}$ is an assumption set w.r.t. M but not an unfounded set, by definition of P-stable model. Therefore, by Lemma 2, M is a fixpoint of W_P. □

Theorem 2 states that no other (positive or negative) fact can be added to a P-stable model using the transformation W_P; so a P-stable model seems to fulfill the principle of minimal undefinedness. But, as pointed out in Example 4, a given program might have several P-stable models, some of which are subsets of others. Thus, P-stable models do not fully respect the principle of minimal undefinedness.

Well-founded models were introduced in [VRS].

Definition 9. Let P be a program. The *well-founded* model of P is the least fixpoint of W_P. □

THEOREM 3. *Let P be a program. The well-founded model of P exists and has the following properties:*

(a) *it is P-stable,*

(b) *it is the intersection of all P-stable models for P*

PROOF. Since W_P is monotonic in a lower semi-lattice, the least fixpoint of W_P exists; so does the well-founded model of P, that will be denoted by M.

(a) Since M is a fixpoint of W_P, by Lemma 1 M is a model of P and every non-empty subset of $\overline{M}$ is an assumption set w.r.t. M but not an unfounded set. Using an argument similar to that of the proof of Lemma 2, it is easy to show that no subset of M^+ is an assumption set w.r.t. M. It follows that M is P-stable.

(b) Given an arbitrary P-stable model N, N is a fixpoint of W_P by Theorem 2 and, then, $M \subseteq N$ for M is the least fixpoint of W_P. It follows that, since M is P-stable, M is the intersection of all P-stable models for P. □

COROLLARY 1. *There exists a P-stable model for every logic program.* □

4. 3-Valued Models

We will next show that the definition of P-stable model (Definition 8) is equivalent to the definition of strongly-founded model [SZ] as well as to that of 3-valued stable model as given in [P3, P4]. To this end, following [P3, P4], we define a *3-valed logic* with T (*true*), F (*false*), and U (*undefined*), ordered as $F < U < T$. Given a program P, an interpretation I, and a ground literal A, *value*$_I(A)$ is T if A is in I, F if $\neg A$ is in I and U if neither A nor $\neg A$ is in I. Moreover, the value of a conjunction C of ground literals is the minimal value of the literals in the conjunction, i.e., $value_I(C) = \min_{A \text{ in } C}(value_I(A))$. If C is empty the we assume that $value_I(C) = T$. Finally, a ground rule r is *satisfied* by I if $value_I(H(r)) \geq value_I(B(r))$.

Definition 10. 3-Valued Model: Let P be a program and I be an interpretation of it. Then I is a *3-valued model* of P if every rule r in $ground(P)$ is satisfied by M. □

We next show that 3-valued models are a subclass of partial models.

LEMMA 3. *Let P be a program and I be an interpretation of it. Then, I is a 3-valued model of P if and only if I is a partial model of P and every non-empty subset of $\overline{I}$ is an assumption set w.r.t. I.*

PROOF. *(If-Part).* Suppose that I is a partial model of P. If $\overline{I}$ is empty then obviously I is 3-valued. Let us then assume that $\overline{I}$ is not empty and that every non-empty subset of $\overline{I}$ is an assumption set w.r.t. I. Let r be an arbitrary rule in $ground(P)$ and $A = H(r)$. If A is in I^+ then obviously $value_I(A) \geq value_I(B(r))$. Moreover, if $\neg A$ is instead in I^-, then, by Part (a) of Proposition 1, the rule is blocked w.r.t. I. Thus, at least one of the goals of r has value F; so $value_I(B(r)) = F$ and, then, $value_I(A) \geq value_I(B(r))$. Finally, suppose that that A is in $\overline{I}$, i.e., $value_I(A) = U$. Since $\{A\}$ is an assumption set w.r.t. I, r is inapplicable w.r.t. I, i.e., $value_I(B(r)) < T$. It follows that $value_I(A) \geq value_I(B(r))$ also in this case and, therefore, I is a 3-valued model of P.

(Only-If-Part). Suppose that I is a 3-valued model of P. Let r be an arbitrary rule in $ground(P)$ such that $\neg A$ is in I^-, where $A = H(r)$. Since $value_I(A) = F$, by Definition 10 $value_I(B(r)) = F$ as well. Hence, at least one of the goals of r has value F, i.e., r is blocked w.r.t. I. Therefore, by Part (a) of Proposition 1, I is a partial model of P. We prove by contradiction that every non-empty subset of $\overline{I}$ is an assumption set w.r.t. I. Suppose then that $X \subseteq \overline{I}$ is not an assumption set w.r.t. I. By definition of assumption set, there exists a rule r in $ground(P)$ with head A such that A is in X, $B(r) \subseteq I$ (i.e., r is not inapplicable w.r.t. I) and none of the goals of r is in X. Hence, $B(r) \subseteq (I - X)$; so, $value_I(B(r)) = T$. Since $value_I(A) = U$, we derive that $value_I(A) < value_I(C)$ — a contradiction with the hypothesis that I is 3-valued. Therefore, every non-empty subset of $\overline{I}$ is an assumption set w.r.t. I and this concludes the proof. □

Example 4. Consider the following program:

$a.$

$b \leftarrow \neg c, c.$

$c \leftarrow b.$

$d \leftarrow e.$

Here we have that the total model $\{a, \neg b, \neg c, \neg d, \neg e\}$ is both stable and 3-valued. The partial model $\{a, c\}$ is 3-valued but not P-stable; the empty set is a partial model but is not 3-valued. □

In [SZ1], strongly-founded models are defined as follows

Definition 11. Strongly-Founded Model: A partial model M is *strongly-founded* if the following three conditions hold:

a) no subset of M^+ is an assumption set w.r.t. M;

b) no subset of $\overline{M}$ is an unfounded set w.r.t. M;

c) M is a 3-valued model for M. □

We next prove that the notions of strongly-founded model and P-stable model coincide.

THEOREM 3. *Definition 11 and Definition 8 are equivalent.*

PROOF. By Lemma 3, a model M is 3-valued if and only if every non-empty subset of $\overline{M}$ is an assumption set w.r.t. M. It follows that Conditions (b) and (c) of Definition 11 can be replaced by Condition (b) of Definition 8, thus the two definitions are equivalent. □

Let us now introduce the notion of 3-valued stable model as given in [P3, P4]. To this end we first introduce some preliminary notations and results.

A *3-valued program* is a program where the constants T, F and U may occur as goals of the rules. Such goals can be thought of as already interpreted ground literals. Obviously, the notion of 3-valued model also holds for 3-valued programs.

Let P be a positive 3-valued program. A 3-valued model M of P is *minimal* if for each 3-valued model M_1 of P, (a) $M^+ \subseteq M_1^+$ and (b) if $M^+ = M_1^+$ then also $M_1^- \subseteq M^-$. In [P3, P4] it has been proved that P has a unique 3-valued minimal model. We next show that this model has an interesting characterization in terms of the notions of immediate consequence transformation T_P of P and of unfounded set. Let us first rephrase the definition of T_P as follows:

$$T_P(I) = \{A \mid r \in ground(P), A = H(r), value_I(B(r)) = T\}$$

where I is an interpretation of P. We also recall that $U_P(I)$ denotes the greatest unfounded set w.r.t. I.

LEMMA 4. *Let P be a positive 3-valued program. Then $T_P^\infty(\emptyset) \cup \neg U_P(T_P^\infty(\emptyset))$ is the unique minimal 3-valued model of P.*

PROOF. Consider the interpretation M, where $M^+ = T_P^\infty(\emptyset)$ and $M^- = \neg U_P(M^+)$. Let r be any rule in $ground(P)$ and $A = H(r)$. If $value_M(A) = T$ then r is trivially satisfied by M. If $value_M(A) = F$, then $value_M(B(r)) = F$ by construction of M and by definition of unfounded set; so, also in this case, r is satisfied by M. Finally, if $value_M(A) = U$ then A is not in M^+; so $value_M(B(r)) < T$ by definition of T_P. It follows that r is satisfied by M also in the last case, thus M is a 3-valued model of P. Let us now prove that, given any 3-valued model N of P, $M^+ = T_P^\infty(\emptyset) \subseteq N$. To this end, as $T_P^\infty(\emptyset) \subseteq T_P^\infty(N)$ because T_P is monotonic, it is sufficient to show that $T_P^\infty(N) \subseteq N$ By definition of least fixpoint, $T_P^\infty(N)$ is equal to $T^k(N)$, for some $k \geq 0$, where $T^0 = N$ and $T^i = T_P(T^{i-1})$, $i \geq 0$. We show that for each i, $0 \leq i \leq k$, $T^i \subseteq N$. We proceed by induction on i. Obviously, $N \subseteq N$; so the basis of the induction trivially holds. Consider now any i, $1 \leq i \leq k$. By inductive hypothesis, $T^{i-1} \subseteq N$. Consider any A in T^i. By definition of T_P, there exists a rule r in *ground*(P) such that $value_{T^{i-1}}(B(r)) = T$. Since $T^{i-1} \subseteq N$, also $value_N(B(r)) = T$. Therefore, $H(r)$ is also in N because otherwise $value_N(H(r)) < value_N(B(r))$ (a contradiction with the hypothesis that N is a 3-valued model). Hence, $T_P^\infty(N) = T^k \subseteq N$ and, then, $M^+ = T_P^\infty(\emptyset) = T^k \subseteq N$. Let N be any 3-valued model of P such that $N^+ = M^+$. Let $X = N^- - M^-$. By definition of 3-valued model, for each A in X and for each rule r in *ground*(P) with $H(r) = A$, $value_N(B(r)) = F$. Hence, there exists at least on goal of r, say B, such that $value_N(B) = F$. There are three possible cases: (1) $B = F$, (2) $\neg B$ is in both M^- and N^-, and (3) B is in X. In the first case, $value_{M^+} = F$, thus r is blocked w.r.t. M^+. In the other two cases, $B(r) \cap (\neg M^- \cup X) \neq \emptyset$. Hence, since $\neg M^-$ is an unfounded set w.r.t. M^+ by construction, $\neg M^- \cup X$ is also an unfounded set w.r.t. M^+. But $\neg M^-$ is the greatest unfounded set w.r.t. M^+; so $X = \emptyset$. It follows that M is the unique minimal 3-valued model of P. □

Let us now consider an ordinary logic program P. Given an interpretation I of P, the *GL-transformation* of P w.r.t. I is the program $GL(P)_I$ obtained from *ground*(P) by replacing in every rule all negative goals $\neg A$ with (a) T if $\neg A$ is in I^-, (b) with F if A is in I, or (c) with U if A is in $\bar{I}$. Obviously $GL(P)_I$ is 3-valued positive program.

Definition 12. 3-Valued Stable Model: A 3-valued model M is *stable* if M is the minimal 3-valued model of $GL(P)_M$. □

We conclude by showing that P-stable models also coincide with 3-valued stable models.

THEOREM 4. *Definitions 12 and 8 are equivalent.*

PROOF. Let P be a program. We first observe that, by Lemma 3, a necessary condition for an interpretation of P to be a P-stable model of P according to both definitions is that it is that it is 3-valued model. Therefore, Fact 2 can be rephrased as follows: "a 3-valued model M of P is P-stable (according to Definition 8) if and only if $T_{P_M}^\infty(\emptyset) = M^+$ and no subset of $\overline{M}$ is an unfounded set w.r.t. M in P". On the other hand, by Lemma 4, Definition 12 can be rephrased as follows: "a 3-valued model M of P is P-stable (according to Definition 12) if and only if $T_{GL(P)}^\infty(\emptyset) = M^+$ and no subset of $\overline{M}$ is an unfounded set w.r.t. M in $GL(P)$". But, considering the definitions of positive version and GL-transformation, it is easy to see that (a) $T_{P_M}^\infty(\emptyset) = T_{GL(P)}^\infty(\emptyset)$ and (b) a subset of $\overline{M}$ is an unfounded set w.r.t. M in P if and only if it is an unfounded set w.r.t. M in $GL(P)$. It follows that Definitions 12 and 8 are equivalent. □

5. Conclusion

It is therefore possible to develop a complete formal theory of partial models using only two-valued logic. In doing so in this paper, we have unveiled important properties of well-founded and stable models, and shown that the parallel interpretation using three-valued logic produces the same results. Therefore, it follows that 3-valued logic is not needed for the treatment of logic programs with negation, although it can be expedient in many situations.

It also follows that the key issue in dealing with logic programs with negation is not the use of two-valued or multi-valued logic, as some authors seem to suggest; rather, the key issue is what design principles should be achieved by the intended semantics.

For instance, while P-stable models enforce the principle of justifiability, they do not obey that of minimal undefinedness, inasmuch as a program can have several P-stable models, some of which are subsets of others. Our claim, therefore, is that the minimal undefinedness principle should be enforced by restricting our attention to the class of P-stable models that are maximal, and that these maximal models should be regarded as the natural generalization of the notion of total stable models to the domain of partial interpretations. A detailed treatment of this approach is presented in [SZ2].

Acknowledgments

The authors thank Els Laenens for her comments on the original manuscript.

References

[ABW] Apt, K., Bair, H., and Walker, A., "*Towards a Theory of Declarative Knowledge*," Minker, J. (ed.), Morgan Kaufman, Los Altos, 1987, pp. 89-148.

[Ceta] Chimenti, et al, "the LDL System Prototype," *IEEE Trans. on Knowledge and Data Engineering*, Vol. 2, No. 1, March 1990.

[CH] Chandra, A., Harel, D., "Horn Clauses and Generalization", *Journal of Logic Programming 2*, 1, 1985, pp. 320-340.

[Do] Doyle, J., A Truth Maintenance System, *Artificial Intelligence*, Vol. 12, (1979), pp. 231-271.

[GL] Gelfond, M., Lifschitz, V., "The Stable Model Semantics for Logic Programming", *Proc. 5th Int. Conf. and Symp. on Logic Programming*, MIT Press, Cambridge, Ma, 1988, pp. 1070-1080.

[GPZ] Giannotti, F., D. Pedreschi and C. Zaniolo, "Declarative Semantics for Pruning Operators in Logic Programming", *Proc. NACLP90 Workshop* on Logic Programming and Non-Monotonic Reasoning, Austin, Tx, Nov 1,2, 1989.

[L] Lloyd, J.W., *Foundations of Logic Programming*, Springer Verlag, Berlin, 1987.

[LV] Laenens, E. and D. Vermeir, "A Fixpoint Semantics of Ordered Logic", *Journal of Logic and Computation*, pp. 159-185. vol. 1, no. 2, 1990.

[MS] Marek, W., and V.S. Subrahmanian, "The relationship between stable, supported defualt, and autoepistemic semantics of general logic programs, Tech. Rep. No. 128-88, University of Kentucky, 1988.

[N] Naqvi, S.A., "A Logic for Negation in Database Systems," in *Foundations of Deductive Databases and Logic Programming, (Minker, J. ed.)*, Morgan Kaufman, Los Altos, 1987.

[P1] Przymusinski, T.C., "On the Declarative and Procedural Semantics of Deductive Databases and Logic Programs", in *Foundations of Deductive Databases and Logic Programming, (Minker, J. ed.)*, Morgan Kaufman, Los Altos, 1987, pp. 193-216.

[P2] Przymusinski, T.C., "Every Logic Program has a Natural Stratification and an Iterated Fixed Point Model", *Proc. ACM Symp. on Principles of Database Systems*, pp. 11-21, 1989.

[P3] Przymusinski, T.C., "Well-Founded Semantics Coincides with Three-Valued Stable-Semantics," *Foundamenta Informaticae*, vol. 13, pp. 445-463, 1990.

[P4] Przymusinski, T.C., "Extended stable semantics for normal and disjunctive programs", *Proc. of the 7th Int. Conf. on Logic Programming*, MIT Press, Cambridge, 1990, pp. 459-477.

[SZ1] Saccà D. and C. Zaniolo, "Stable Models and Non-Determinism in Logic Programs with Negation," *Proc. ACM PODS Symp.*, 1990, pp.205-217.

[SZ2] Saccà D. and C. Zaniolo, "Stable Models in Logic Programs with Negation and Non-Determinism", MCC Technical Report, 1991.

[U] Ullman, J.D., *Principles of Database and Knowledge-Base Systems, Vol. 1 and 2*, Computer Science Press, Rockville, Md., 1989.

[V1] Van Gelder, A., "Negation as Failure Using Tight Derivations for Logic Programs," *Proc. 3rd IEEE Symp. on Logic Programming*, Springer-Verlag, 1986, pp. 127-138.

[V2] Van Gelder, A., "The Alternating Fixpoint of Logic Programs with Negation", *Proc. ACM Symp. on Principles of Database Systems*, pp. 1-10, 1989.

[VRS] Van Gelder, A., Ross, K., Schlipf, J.S., "Unfounded Sets and Well-Founded Semantics for General Logic Programs", *ACM SIGMOD-SIGACT Symp. on Principles of Database Systems*, March 1988, pp. 221-230.

[YoYu] You,J-H. and L.Y. Yuan, "Three-Valued Formalization in Logic Programs: is it really needed?", *ACM SIGMOD-SIGACT Symp. on Principles of Database Systems*, pp. 172-181, 1990.