Circular Belief in Logic Programming Semantics

Pascal Hitzler

Knowledge Representation and Reasoning Group, Artificial Intelligence Institute
Department of Computer Science, Dresden University of Technology
Dresden, Germany

phitzler@inf.tu-dresden.de
www.wv.inf.tu-dresden.de/~pascal/

Technical Report WV–02–13

Abstract

In [HW02b, HW02a], a new methodology has been proposed which allows to derive uniform characterizations of different declarative semantics for logic programs with negation, and it has also been hinted at the possibility to use this novel approach in order to obtain new meaningful semantics for logic programs. In this paper, we substantiate this claim by proposing a new semantics which allows to deal with circular belief in logic programming. We will also show that this *circular semantics* makes it possible to encode uncertain knowledge with logic programs in a novel way.

Contents

1 Introduction

Negation in logic programming differs from the negation of classical logic. Indeed, the quest for a satisfactory understanding of negation in logic programming is still inconclusive — although the issue has cooled down a bit recently — and has proved to be very stimulating for research activities in computational logic, and in particular amongst knowledge representation and reasoning researchers concerned with commonsense and nonmonotonic reasoning. During the last two decades, different interpretations of negation in logic programming have lead to the development of a variety of *declarative semantics*, as they are called. Some early research efforts for establishing a satisfactory declarative semantics for negation as failure and its variants, as featured by the resolution-based Prolog family of logic programming systems, have later on been merged with nonmonotonic frameworks for commonsense reasoning, culminating recently in the development of so-called answer set programming systems, like SMODELS or DLV [MT99, SNS0x]. In turn, these have recently been shown to have potential for representing action and change in research on cognitive agents [Lif02].

Systematically, one can understand Fitting's proposal [Fit85] of a *Kripke-Kleene semantics* — also known as *Fitting semantics* — as a cornerstone which plays a fundamental role both for resolution-based and nonmonotonic reasoning inspired logic programming. Indeed, his proposal, which is based on a monotonic semantic operator in Kleene's strong three-valued logic, has been pursued in both communities, for example by Kunen [Kun87] for giving a semantics for pure Prolog, and by Apt and Pedreschi [AP93] in their fundamental paper on termination analysis of negation as failure, leading to the notion of *acceptable program*. On the other hand, however, Fitting himself [Fit91, Fit02], using a bilattice-based approach which was further developed by Denecker, Marek and Truszczynski [DMT00], tied his semantics closely to the major semantics inspired by nonmonotonic reasoning, namely the *stable model semantics* due to Gelfond and Lifschitz [GL88], which is based on a nonmonotonic semantic operator, and the *well-founded semantics* due to van Gelder, Ross and Schlipf [vGRS91], originally defined using a different monotonic operator in three-valued logic together with a notion of unfoundedness.

Another fundamental idea which was recognised in both communities was that of *stratification*, with the underlying idea of restricting attention to certain kinds of programs in which recursion through negation is prevented. Apt, Blair and Walker [ABW88] proposed a variant of resolution suitable for these programs, while Przymusinski [Prz88] coined the slighty more general notion of *local stratification*.

The semantics mentioned so far are defined and characterized using a variety of different techniques and constructions, including monotonic and nonmonotonic semantic operators in two- and three-valued logics, program transformations, level mappings, restrictions to suitable subprograms, detection of cyclic dependencies etc. So in [HW02b, HW02a], the authors have proposed a methodology which allows to obtain uniform characterizations of all semantics previously mentioned, and which clearly has potential for encompassing more than these. The characterizations allow immediate comparison between the semantics, and one interesting observation made in [HW02b, HW02a] was the fact that the well-founded semantics can formally be understood as a Fitting semantics augmented with a form of stratification. We will reproduce this argument at the end of Section 2 since it is important for the motivation underlying our new proposal. In fact, while this perspective is very appealing, it suffers from a certain asymmetry. Formally, this asymmetry is easily rectified, and a new semantics is born. The remainder of the paper will then substantiate the claim that this new semantics is both satisfactory from a theoretical perspective and meaningful from a knowledge representation point of view. Indeed, on the theoretical side it will turn out to be tightly coupled to variants of the Gelfond-Lifschitz operator, and thus to variants of the stable model semantics. On the applied side it will turn out to be a reasonable framework for dealing with circular, i.e. unfounded, belief. A slight reinterpretation will also make it applicable to reasoning with uncertain knowledge.

The plan of the paper is as follows. Section 2 contains preliminaries which are needed to make the paper relatively self-contained, as well as a short survey of some of the main results from [HW02b, HW02a]. In Section 3 we will use the results from [HW02b, HW02a] just mentioned in order to systematically create a proposal for a new semantics, based on purely formal considerations. In Sections 4 and 5 we will show that the proposed semantics is meaningful, and is indeed a variant of the well-founded semantics, in a strictly formal sense. In Section 6, we will eventually merge the new proposal with the well-founded semantics, in order to arrive at the *circular semantics* for logic programs which we propose for dealing with circular belief. Some examples will illustrate its potential for reasoning with uncertainty. We conclude and discuss possibilities for further work in Section 7.

Acknowledgement. Thanks to Herbert Stoyan for his encouragment to extend [HW02b], and to Matthias Wendt for inspiring discussions. Thanks also to Mike Stange for some terminological first aid.

2 Preliminaries and Notation

A (*normal*) *logic program* is a finite set of (universally quantified) *clauses* of the form $\forall(A \leftarrow A_1 \wedge \cdots \wedge A_n \wedge \neg B_1 \wedge \cdots \wedge \neg B_m)$, commonly written as $A \leftarrow A_1, \ldots, A_n, \neg B_1, \ldots, \neg B_m$, where A, A_i, and B_j, for $i = 1, \ldots, n$ and $j = 1, \ldots, m$, are atoms over some given first order language. A is called the *head* of the clause, while the remaining atoms make up the *body* of the clause, and depending on context, a body of a clause will be a set of literals (i.e. atoms or negated atoms) or the conjunction of these literals. Care will be taken that this identification does not cause confusion. We allow a body, i.e. a conjunction, to be empty, in which case it always evaluates to true. A clause with empty body is called a *unit clause* or a *fact*. A clause is called *definite*, if it contains no negation symbol. A program is called *definite* if it consists only of definite clauses. We will usually denote atoms with A or B, and literals, which may be atoms or negated atoms, by L or K.

Given a logic program P, we can extract from it the components of a first order language, and we always make the mild assumption that this language contains at least one constant symbol. The corresponding set of ground atoms, i.e. the *Herbrand base* of the program, will be denoted by B_P. For a subset $I \subseteq B_P$, we set $\neg I = \{\neg A \mid A \in B_P\}$. The set of all ground instances of P with respect to B_P will be denoted by $\mathsf{ground}(P)$. For $I \subseteq B_P \cup \neg B_P$, we say that A is *true with respect to* (or *in*) I if $A \in I$, we say that A is *false with respect to* (or *in*) I if $\neg A \in I$, and if neither is the case, we say that A is *undefined with respect to* (or *in*) I. A (*three-valued* or *partial*) *interpretation* I for P is a subset of $B_P \cup \neg B_P$ which is *consistent*, i.e.

whenever $A \in I$ then $\neg A \notin I$. A body, i.e. a conjunction of literals, is true in an interpretation I if every literal in the body is true in I, it is false in I if one of its literals is false in I, and otherwise it is undefined in I. For a negated literal $L = \neg A$ we will find it convenient to write $\neg L \in I$ if $A \in I$. By I_P we denote the set of all (three-valued) interpretations of P. Both I_P and $B_P \cup \neg B_P$ are complete partial orders (cpos) via set-inclusion, i.e. they contain the empty set as least element, and every ascending chain has a supremum, namely its union. A *model* of P is an interpretation $I \in I_P$ such that for each clause $A \leftarrow \mathsf{body}$ we have that $\mathsf{body} \subseteq I$ implies $A \in I$. A *total* interpretation is an interpretation I such that no $A \in B_P$ is undefined in I.

For an interpretation I and a program P, an *I-partial level mapping* for P is a partial mapping $l : B_P \to \alpha$ with domain $\mathsf{dom}(l) = \{A \mid A \in I \text{ or } \neg A \in I\}$, where α is some (countable) ordinal. We extend every level mapping to literals by setting $l(\neg A) = l(A)$ for all $A \in \mathsf{dom}(l)$. A (*total*) *level mapping* is a total mapping $l : B_P \to \alpha$ for some (countable) ordinal α.

Given a normal logic program P and some $I \subseteq B_P \cup \neg B_P$, we say that $U \subseteq B_P$ is an *unfounded set* (*of* P) *with respect to* I if each atom $A \in U$ satisfies the following condition: For each clause $A \leftarrow \mathsf{body}$ in $\mathsf{ground}(P)$ (at least) one of the following holds.

(Ui) Some (positive or negative) literal in body is false in I.

(Uii) Some (non-negated) atom in body occurs in U.

Given a normal logic program P, we define the following operators on $B_P \cup \neg B_P$. $T_P(I)$ is the set of all $A \in B_P$ such that there exists a clause $A \leftarrow \mathsf{body}$ in $\mathsf{ground}(P)$ such that body is true in I. $F_P(I)$ is the set of all $A \in B_P$ such that for all clauses $A \leftarrow \mathsf{body}$ in $\mathsf{ground}(P)$ we have that body is false in I. Both T_P and F_P map elements of I_P to elements of I_P. Now define the operator $\Phi_P : I_P \to I_P$ by

$$\Phi_P(I) = T_P(I) \cup \neg F_P(I).$$

This operator is due to [Fit85] and is well-defined and monotonic on the cpo I_P, hence has a least fixed point by the Knaster-Tarski[1] fixed-point theorem, and we can obtain this fixed point by defining, for each monotonic operator F, that $F\uparrow 0 = \emptyset$, $F\uparrow(\alpha+1) = F(F\uparrow\alpha)$ for any ordinal α, and $F\uparrow\beta = \bigcup_{\gamma<\beta} F\uparrow\gamma$ for any limit ordinal β, and the least fixed point of F is obtained as $F\uparrow\alpha$ for some ordinal α. The least fixed point of Φ_P is called the *Kripke-Kleene model* or *Fitting model* of P, determining the *Fitting semantics* of P.

Now, for $I \subseteq B_P \cup \neg B_P$, let $U_P(I)$ be the greatest unfounded set (of P) with respect to I, which always exists due to [vGRS91]. Finally, define

$$W_P(I) = T_P(I) \cup \neg U_P(I)$$

[1]We follow the terminology from [Jac01]. The Knaster-Tarski theorem is sometimes called Tarski theorem and states that every monotonic function on a cpo has a least fixed point, which can be obtained by transfinitely iterating the bottom element of the cpo. The Tarski-Kantorovitch theorem is sometimes refered to as the Kleene theorem or the Scott theorem (or even as "the" fixed-point theorem) and states that if the function is additionally Scott (or order-) continuous, then the least fixed point can be obtained by an iteration which is not transfinite, i.e. closes off at ω, the least infinite ordinal. In both cases, the least fixed point is also the least pre-fixed point of the function.

for all $I \subseteq B_P \cup \neg B_P$. The operator W_P, which operates on the cpo $B_P \cup \neg B_P$, is due to [vGRS91] and is monotonic, hence has a least fixed point by the Tarski fixed-point theorem, as above for Φ_P. It turns out that $W_P \uparrow \alpha$ is in I_P for each ordinal α, and so the least fixed point of W_P is also in I_P and is called the *well-founded model* of P, giving the *well-founded semantics* of P.

In order to avoid confusion, we will use the following terminology: the notion of *interpretation*, and I_P will be the set of all those, will by default denote consistent subsets of $B_P \cup \neg B_P$, i.e. interpretations in three-valued logic. We will sometimes emphasize this point by using the notion *partial interpretation*. By *two-valued interpretations* we mean subsets of B_P. Both interpretations and two-valued interpretations are ordered by subset inclusion. Each two-valued interpretation I can be identified with the partial interpretation $I' = I \cup \neg(B_P \setminus I)$. Note however, that in this case I' is always a maximal element in the ordering for partial interpretations, while I is in general not maximal as a two-valued interpretation[2]. Given a partial interpretation I, we set $I^+ = I \cap B_P$ and $I^- = \{A \in B_P \mid \neg A \in I\}$.

Given a program P, we define the operator T_P^+ on subsets of B_P by $T_P^+(I) = T_P(I \cup \neg(B_P \setminus I))$. The pre-fixed points of T_P^+, i.e. the two-valued interpretations $I \subseteq B_P$ with $T_P^+(I) \subseteq I$, are exactly the models, in the sense of classical logic, of P. Post-fixed points of T_P^+, i.e. $I \subseteq B_P$ with $I \subseteq T_P^+(I)$ are called *supported interpretations* of P, and a supported model of P is a model P which is a supported interpretation. The supported models of P thus coincide with the fixed points of T_P^+. It is well-known that for definite programs P the operator T_P^+ is monotonic on the set of all subsets of B_P, with respect to subset inclusion. Indeed it is Scott-continuous [Llo88, SHLG94] and, via the Tarski-Kantorovich[1] fixed-point theorem, achieves its least pre-fixed point M, which is also a fixed point, as the supremum of the iterates $T_P^+ \uparrow n$ for $n \in \mathbb{N}$. So $M = \mathsf{lfp}\left(T_P^+\right) = T_P^+ \uparrow \omega$ is *the least two-valued model* of P. Likewise, since the set of all subsets of B_P is a complete lattice, and therefore has greatest element B_P, we can also define $T_P^+ \downarrow 0 = B_P$ and inductively $T_P^+ \downarrow (\alpha + 1) = T_P^+(T_P^+ \downarrow \alpha)$ for each ordinal α and $T_P^+ \downarrow \beta = \bigcap_{\gamma<\beta} T_P^+ \downarrow \gamma$ for each limit ordinal β. Again by the Knaster-Tarski fixed-point theorem, applied to the superset inclusion ordering (i.e. reverse subset inclusion) on subsets of B_P, it turns out that T_P^+ has a greatest fixed point, $\mathsf{gfp}\left(T_P^+\right)$.

There is a semantics using two-valued logic, the stable model semantics due to [GL88], which is intimately related to the well-founded semantics. Let P be a normal program, and let $M \subseteq B_P$ be a set of atoms. Then we define P/M to be the (ground) program consisting of all clauses $A \leftarrow A_1, \ldots, A_n$ for which there is a clause $A \leftarrow A_1, \ldots, A_n, \neg B_1, \ldots, \neg B_m$ in $\mathsf{ground}(P)$ with $B_1, \ldots, B_m \notin M$. Since P/M does no longer contain negation, it has a least two-valued model $T_{P/M}^+ \uparrow \omega$. For any two-valued interpretation I we can therefore define the operator $\mathrm{GL}_P(I) = T_{P/I}^+ \uparrow \omega$, and call M a *stable model* of the normal program P if it is a fixed point of the operator GL_P, i.e. if $M = \mathrm{GL}_P(M) = T_{P/M}^+ \uparrow \omega$. As it turns out, the operator GL_P is in general not monotonic for normal programs P. However it is *antitonic*, i.e. whenever $I \subseteq J \subseteq B_P$ then $\mathrm{GL}_P(J) \subseteq \mathrm{GL}_P(I)$. As a consequence, the operator GL_P^2, obtained by applying GL_P twice, is monotonic, and hence has a least fixed point L_P and a greatest fixed point G_P. In [vG89] it was shown that $\mathrm{GL}_P(L_P) = G_P$, $L_P = \mathrm{GL}_P(G_P)$, and that $L_P \cup \neg(B_P \setminus G_P)$ coincides with the well-founded model of P. This is called the *alternating fixed point characterization* of the well-founded semantics.

[2]These two orderings in fact correspond to the knowledge and truth orderings as discussed in [Fit91].

Some Results

The following is a straightforward result which has, to the best of our knowledge, not been noted before. It follows the general approach put forward in [HW02b, HW02a].

2.1 Theorem Let P be a definite program. Then there is a unique two-valued model M of P for which there exists a (total) level mapping $l : B_P \to \alpha$ such that for each atom $A \in M$ there exists a clause $A \leftarrow A_1, \ldots, A_n$ in $\mathsf{ground}(P)$ with $A_i \in M$ and $l(A) > l(A_i)$ for all $i = 1, \ldots, n$. Furthermore, M is the least two-valued model of P.

Proof: Let M be the least two-valued model $T_P^+\!\uparrow\!\omega$, choose $\alpha = \omega$, and define $l : B_P \to \alpha$ by setting $l(A) = \mathsf{min}\{n \mid A \in T_P^+\!\uparrow\!(n+1)\}$, if $A \in M$, and by setting $l(A) = 0$, if $A \notin M$. From the fact that $\emptyset \subseteq T_P^+\!\uparrow\!1 \subseteq \ldots \subseteq T_P^+\!\uparrow\!n \subseteq \ldots \subseteq T_P^+\!\uparrow\!\omega = \bigcup_m T_P^+\!\uparrow\!m$, for each n, we see that l is well-defined and that the least model $T_P^+\!\uparrow\!\omega$ for P has the desired properties.

Conversely, if M is a two-valued model for P which satisfies the given condition for some mapping $l : B_P \to \alpha$, then it is easy to show, by induction on $l(A)$, that $A \in M$ implies $A \in T_P^+\!\uparrow\!(l(A)+1)$. This yields that $M \subseteq T_P^+\!\uparrow\!\omega$, and hence that $M = T_P^+\!\uparrow\!\omega$ by minimality of the model $T_P^+\!\uparrow\!\omega$. ■

The following result is due to [Fag94], and is striking in its similarity to Theorem 2.1.

2.2 Theorem Let P be normal. Then a two-valued model $M \subseteq B_P$ of P is a stable model of P if and only if there exists a (total) level mapping $l : B_P \to \alpha$ such that for each $A \in M$ there exists $A \leftarrow A_1, \ldots, A_n \neg B_1, \ldots, \neg B_m$ in $\mathsf{ground}(P)$ with $A_i \in M$, $B_j \notin M$, and $l(A) > l(A_i)$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

We next recall the following alternative characterization of the Fitting model, due to [HW02b, HW02a].

2.3 Definition Let P be a normal logic program, I be a model of P, and l be an I-partial level mapping for P. We say that P *satisfies* (F) *with respect to* I *and* l, if each $A \in \mathsf{dom}(l)$ satisfies one of the following conditions.

(Fi) $A \in I$ and there exists a clause $A \leftarrow L_1, \ldots, L_n$ in $\mathsf{ground}(P)$ such that $L_i \in I$ and $l(A) > l(L_i)$ for all i.

(Fii) $\neg A \in I$ and for each clause $A \leftarrow L_1, \ldots, L_n$ in $\mathsf{ground}(P)$ there exists i with $\neg L_i \in I$ and $l(A) > l(L_i)$.

If $A \in \mathsf{dom}(l)$ satisfies (Fi), then we say that A *satisfies* (Fi) *with respect to* I *and* l, and similarly if $A \in \mathsf{dom}(l)$ satisfies (Fii).

2.4 Theorem Let P be a normal logic program with Fitting model M. Then M is the greatest model among all models I, for which there exists an I-partial level mapping l for P such that P satisfies (F) with respect to I and l.

Let us recall next the definition of a (locally) stratified program, due to [ABW88, Prz88]: A normal logic program is called *locally stratified* if there exists a (total) level mapping

$l : B_P \to \alpha$, for some ordinal α, such that for each clause $A \leftarrow A_1, \ldots, A_n, \neg B_1, \ldots, \neg B_m$ in $\mathsf{ground}(P)$ we have that $l(A) \geq l(A_i)$ and $l(A) > l(B_j)$ for all $i = 1, \ldots, n$ and $j = 1, \ldots, m$. The notion of (locally) stratifed program was developed with the idea of preventing *recursion through negation*, while allowing recursion through positive dependencies. (Locally) stratified programs have total well-founded models.

There exist locally stratified programs which do not have a total Fitting semantics and vice versa. In fact, condition (Fii) requires a strict decrease of level between the head and a literal in the rule, independent of this literal being positive or negative. But, on the other hand, condition (Fii) imposes no further restrictions on the remaining body literals, while the notion of local stratification does. These considerations motivate the substitution of condition (Fii) by the condition (Cii), as done for the following definition.

2.5 Definition Let P be a normal logic program, I be a model of P, and l be an I-partial level mapping for P. We say that P *satisfies* (WF) *with respect to* I *and* l, if each $A \in \mathsf{dom}(l)$ satisfies one of the following conditions.

(Fi) $A \in I$ and there exists a clause $A \leftarrow L_1, \ldots, L_n$ in $\mathsf{ground}(P)$ with $L_i \in I$ and $l(A) > l(L_i)$ for all i.

(Cii) $\neg A \in I$ and for each clause $A \leftarrow A_1, \ldots, A_n, \neg B_1, \ldots, \neg B_m$ in $\mathsf{ground}(P)$ (at least) one of the following conditions holds:

- (Ciia) There exists $i \in \{1, \ldots, n\}$ with $\neg A_i \in I$ and $l(A) \geq l(A_i)$.
- (Ciib) There exists $j \in \{1, \ldots, m\}$ with $B_j \in I$ and $l(A) > l(B_j)$.

If $A \in \mathsf{dom}(l)$ satisfies (Fi), then we say that A *satisfies* (Fi) *with respect to* I *and* l, and similarly if $A \in \mathsf{dom}(l)$ satisfies (Cii).

So, in the light of Theorem 2.4, Definition 2.5 should provide a natural "stratified version" of the Fitting semantics. And indeed it does, and furthermore, the resulting semantics coincides with the well-founded semantics, which is a very satisfactory result from [HW02b, HW02a].

2.6 Theorem Let P be a normal logic program with well-founded model M. Then M is the greatest model among all models I, for which there exists an I-partial level mapping l for P such that P satisfies (WF) with respect to I and l.

For completeness, we remark that an alternative characterization of the weakly perfect model semantics [PP90] can also be found in [HW02b, HW02a].

The approach which led to the results just mentioned, originally put forward in [HW02b, HW02a], provides a methodology for obtaining uniform characterizations of different semantics for logic programs. It appears to be obvious that by further modifying the conditions in Definitions 2.3 or 2.5, one can obtain new semantics, in a formal way. Certainly, such an undertaking will be of purely academic value unless the new semantics can be shown, at hindsight, to represent something meaningful. That is exactly what we will undertake in the following.

3 A Proposal for a New Semantics

We note that condition (Fi) has been reused in Definition 2.5. Thus, Definition 2.3 has been "stratified" only with respect to condition (Fii), yielding (Cii), but not with respect to (Fi). Indeed, also replacing (Fi) by a stratified version such as the following seems not satisfactory at first sight.

(Ci) $A \in I$ and there exists a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in $\mathsf{ground}(P)$ such that $A_i, \neg B_j \in I$, $l(A) \geq l(A_i)$, and $l(A) > l(B_j)$ for all i and j.

If we replace condition (Fi) by condition (Ci) in Definition 2.5, then it is not guaranteed that for any given program there is a greatest model satisfying the desired properties, as the following example from [HW02b, HW02a] shows.

3.1 Example Consider the program consisting of the two clauses $p \leftarrow p$ and $q \leftarrow \neg p$, and the two (total) models $M_1 = \{p, \neg q\}$ and $M_2 = \{\neg p, q\}$, which are incomparable, and the level mapping l with $l(p) = 0$ and $l(q) = 1$.

Nevertheless, we will argue later in Section 6, that conditions (Ci) and (Cii) together provide a meaningful semantics which is suitable for dealing with circular belief and uncertainty. In order to arrive at an understanding of this, we first consider the setting with conditions (Ci) and (Fii), which is somehow "dual" to the well-founded semantics which is characterized by (Fi) and (Cii).

3.2 Definition Let P be a normal logic program, I be a model of P, and l be an I-partial level mapping for P. We say that P *satisfies* (CW) *with respect to* I *and* l, if each $A \in \mathsf{dom}(l)$ satisfies one of the following conditions.

(Ci) $A \in I$ and there exists a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in $\mathsf{ground}(P)$ such that $A_i, \neg B_j \in I$, $l(A) \geq l(A_i)$, and $l(A) > l(B_j)$ for all i and j.

(Fii) $\neg A \in I$ and for each clause $A \leftarrow L_1, \dots, L_n$ in $\mathsf{ground}(P)$ there exists i with $\neg L_i \in I$ and $l(A) > l(L_i)$.

If $A \in \mathsf{dom}(l)$ satisfies (Ci), then we say that A *satisfies* (Ci) *with respect to* I *and* l, and similarly if $A \in \mathsf{dom}(l)$ satisfies (Fii).

In the following two sections, we will present arguments which substantiate the claim that Definition 3.2 provides a meaningful variant of the well-founded semantics. In order to do this, we will start our investigations with a variant of the stable semantics. Via an alternating fixed point characterization, this will also yield a variant of the well-founded semantics, which in turn we will see to be characterized by Definition 3.2.

4 Maximally Circular Stable Semantics

The following result characterizes the greatest model of a definite program analogously to Theorem 2.1.

4.1 Theorem Let P be a definite program. Then there is a unique two-valued supported interpretation M of P for which there exists a (total) level mapping $l : B_P \to \alpha$ such that for each atom $A \notin M$ and for all clauses $A \leftarrow A_1, \ldots, A_n$ in $\mathsf{ground}(P)$ there is some $A_i \notin M$ with $l(A) > l(A_i)$. Furthermore, M is the greatest two-valued model of P.

Proof: Let M be the greatest two-valued model of P, and let α be the least ordinal such that $M = T_P^+ \downarrow \alpha$. Define $l : B_P \to \alpha$ by setting $l(A) = \min\{\gamma \mid A \notin T_P^+ \downarrow (\gamma + 1)\}$ for $A \notin M$, and by setting $l(A) = 0$ if $A \in M$. The mapping l is well-defined because $A \notin M$ with $A \notin T_P^+ \downarrow \gamma = \bigcap_{\beta<\gamma} T_P^+ \downarrow \beta$ for some limit ordinal γ implies $A \notin T_P^+ \downarrow \beta$ for some $\beta < \gamma$. So the least ordinal β with $A \notin T_P^+ \downarrow \beta$ is always a successor ordinal. Now assume that there is $A \notin M$ which does not satisfy the stated condition. We can furthermore assume without loss of generality that A is chosen with this property such that $l(A)$ is minimal. Let $A \leftarrow A_1, \ldots, A_n$ be a clause in $\mathsf{ground}(P)$. Since $A \notin T_P^+ \left(T_P^+ \downarrow l(A)\right)$ we obtain $A_i \notin T_P^+ \downarrow l(A) \supseteq M$ for some i. But then $l(A_i) < l(A)$ which contradicts minimality of $l(A)$.

Conversely, let M be a two-valued model for P which satisfies the given condition for some mapping $l : B_P \to \alpha$. We show by transfinite induction on $l(A)$ that $A \notin M$ implies $A \notin T_P^+ \downarrow (l(A) + 1)$, which suffices because it implies that for the greatest two-valued model $T_P^+ \downarrow \beta$ of P we have that $T_P^+ \downarrow \beta \subseteq M$, and therefore $T_P^+ \downarrow \beta = M$. For the inductive proof consider first the case where $l(A) = 0$. Then there is no clause in $\mathsf{ground}(P)$ with head A and consequently $A \notin T_P^+ \downarrow 1 = T_P^+(B_P)$. Now assume that the statement to be proven holds for all $B \notin M$ with $l(B) < \alpha$, where α is some ordinal, and let $A \notin M$ with $l(A) = \alpha$. Then each clause in $\mathsf{ground}(P)$ with head A contains an atom B with $l(B) = \beta < \alpha$ and $B \notin M$. Hence $B \notin T_P^+ \downarrow (\beta + 1)$ and consequently $A \notin T_P^+ \downarrow (\alpha + 1)$. ■

The following definition and theorem are analogous to Theorem 2.2.

4.2 Definition Let P be normal. Then $M \subseteq B_P$ is called a *maximally circular stable model* (*maxstable model*) of P if it is a two-valued supported interpretation of P and there exists a (total) level mapping $l : B_P \to \alpha$ such that for each atom $A \notin M$ and for all clauses $A \leftarrow A_1, \ldots, A_n, \neg B_1, \ldots, \neg B_m$ in $\mathsf{ground}(P)$ with $B_1, \ldots, B_m \notin M$ there is some $A_i \notin M$ with $l(A) > l(A_i)$.

4.3 Lemma Every maxstable model is a supported model.

Proof: Supportedness follows immediately from the definition. Now assume that M is maxstable but is not a model, i.e. there is $A \notin M$ but there is a clause $A \leftarrow A_1, \ldots, A_n$ in $\mathsf{ground}(P)$ with $A_i \in M$ for all i. But by the definition of maxstable model we must have that there is $A_i \notin M$, which contradicts $A_i \in M$. ■

4.4 Theorem $M \subseteq B_P$ is a maxstable model of P if and only if $M = \mathsf{gfp}\left(T_{P/M}^+\right)$.

Proof: Let M be a maxstable model of P. Let $A \notin M$ and let $T_{P/M}^+ \downarrow \alpha = \mathsf{gfp}\left(T_{P/M}^+\right)$. We show by transfinite induction on $l(A)$ that $A \notin T_{P/M}^+ \downarrow (l(A) + 1)$ and hence $A \notin T_{P/M}^+ \downarrow \alpha$. For $l(A) = 0$ there is no clause with head A in P/M, so $A \notin T_{P/M}^+ \downarrow 1$. Now let $l(A) = \beta$ for some ordinal β. By assumption we have that for all clauses $A \leftarrow A_1, \ldots, A_n, \neg B_1, \ldots, \neg B_m$

with $B_1, \ldots, B_m \notin M$ there exists $A_i \notin M$ with $l(A) > l(A_i)$, say $l(A_i) = \gamma < \beta$. Hence $A_i \notin T^+_{P/M}{\downarrow}(\gamma+1)$, and consequently $A \notin T^+_{P/M}{\downarrow}(\beta+1)$, which shows that $\mathsf{gfp}\left(T^+_{P/M}\right) \subseteq M$.

So let again M be a maxstable model of P and let $A \notin \mathsf{gfp}\left(T^+_{P/M}\right) = T^+_{P/M}{\downarrow}\alpha$ and $l(A) = \beta$. Then for each clause $A \leftarrow A_1, \ldots, A_n$ in P/M there is A_i with $A_i \notin T^+_{P/M}{\downarrow}\alpha$ and $l(A) > l(A_i)$. Now assume $A \in M$. Without loss of generality we can furthermore assume that A is chosen such that $l(A) = \beta$ is minimal. Hence $A_i \notin M$, and we obtain that for each clause in P/M with head A one of the corresponding body atoms is false in M. By supportedness of M this yields $A \notin M$, which contradicts our assumption. Hence $A \notin M$ as desired.

Conversely, let $M = \mathsf{gfp}\left(T^+_{P/M}\right)$. Then as an immediate consequence of Theorem 4.1 we obtain that M is maxstable. ■

5 Maximally Circular Well-Founded Semantics

Maxstable models are formally analogous[3] to stable models in that the former are fixed points of the operator $I \mapsto \mathsf{gfp}\left(T^+_{P/I}\right)$, while the latter are fixed points of the operator $I \mapsto \mathsf{lfp}\left(T^+_{P/I}\right)$. Further, in analogy to the alternating fixed point characterization of the well-founded model, we can obtain a corresponding variant of the well-founded semantics, which we will do next. Theorem 4.4 suggests the defininition of the following operator.

5.1 Definition Let P be a normal program and I be a two-valued interpretation. Then define $\mathrm{CGL}_P(I) = \mathsf{gfp}\left(T^+_{P/I}\right)$.

Using the operator CGL_P, we can define a "maximally circular" version of the alternating fixed-point semantics.

5.2 Proposition Let P be a normal program. Then the following hold.

(i) CGL_P is antitonic and CGL_P^2 is monotonic.

(ii) $\mathrm{CGL}_P\left(\mathsf{lfp}\left(\mathrm{CGL}_P^2\right)\right) = \mathsf{gfp}\left(\mathrm{CGL}_P^2\right)$ and $\mathsf{lfp}\left(\mathrm{CGL}_P^2\right) = \mathrm{CGL}_P\left(\mathsf{gfp}\left(\mathrm{CGL}_P^2\right)\right)$.

Proof: (i) If $I \subseteq J \in B_P$, then $P/J \subseteq P/I$ and consequently $\mathrm{CGL}_P(J) = \mathsf{gfp}\left(T^+_{P/J}\right) \subseteq \mathsf{gfp}\left(T^+_{P/I}\right) = \mathrm{CGL}_P(I)$. Monotonicity of CGL_P^2 then follows trivially.

(ii) Let $L_P = \mathsf{lfp}\left(\mathrm{CGL}_P^2\right)$ and $G_P = \mathsf{gfp}\left(\mathrm{CGL}_P^2\right)$. Then we obtain $\mathrm{CGL}_P^2(\mathrm{CGL}_P(L_P)) = \mathrm{CGL}_P\left(\mathrm{CGL}_P^2(L_P)\right) = \mathrm{CGL}_P(L_P)$, so $\mathrm{CGL}_P(L_P)$ is a fixed point of CGL_P^2, and hence $L_P \subseteq \mathrm{CGL}_P(L_P) \subseteq G_P$. Similarly, $L_P \subseteq \mathrm{CGL}_P(G_P) \subseteq G_P$. Since $L_P \subseteq G_P$ we get from the antitonicity of CGL_P that $L_P \subseteq \mathrm{CGL}_P(G_P) \subseteq \mathrm{CGL}_P(L_P) \subseteq G_P$. Similarly, since $\mathrm{CGL}_P(L_P) \subseteq G_P$, we obtain $\mathrm{CGL}_P(G_P) \subseteq \mathrm{CGL}_P^2(L_P) = L_P \subseteq \mathrm{CGL}_P(G_P)$, so $\mathrm{CGL}_P(G_P) = L_P$, and also $G_P = \mathrm{CGL}_P^2(G_P) = \mathrm{CGL}_P(L_P)$. ■

[3]The term *dual* seems not to be entirely adequate in this situation, although it is intuitionally appealing.

We will now define an operator for the maximally circular well-founded semantics. Given a normal logic program P and some $I \in I_P$, we say that $S \subseteq B_P$ is a *self-founded set* (*of* P) *with respect to* I if $S \cup I \in I_P$ and each atom $A \in S$ satisfies the following condition: There exists a clause $A \leftarrow \texttt{body}$ in $\texttt{ground}(P)$ such that one of the following holds.

(Si) $\texttt{body}$ is true in I.

(Sii) Some (non-negated) atoms in $\texttt{body}$ occur in S and all other literals in $\texttt{body}$ are true in I.

Self-founded sets are analogous[4] to unfounded sets, and the following proposition holds.

5.3 Proposition Let P be a normal program and let $I \in I_P$. Then there exists a greatest self-founded set of P with respect to I.

Proof: If $(S_i)_{i \in \mathcal{I}}$ is a family of sets each of which is a self-founded set of P with respect to I, then it is easy to see that $\bigcup_{i \in \mathcal{I}} S_i$ is also a self-founded set of P with respect to I. ■

Given a normal program P and $I \in I_P$, let $S_P(I)$ be the greatest self-founded set of P with respect to I, and define the operator CW_P on I_P by

$$\mathrm{CW}_P(I) = S_P(I) \cup \neg F_P(I).$$

5.4 Proposition The operator CW_P is well-defined and monotonic.

Proof: For well-definedness, we have to show that $S_P(I) \cap F_P(I) = \emptyset$ for all $I \in I_P$. So assume there is $A \in S_P(I) \cap F_P(I)$. From $A \in F_P(I)$ we obtain that for each clause with head A there is a corresponding body literal L which is false in I. From $A \in S_P(I)$, more precisely from (Sii), we can furthermore conclude that L is an atom and $L \in S_P(I)$. But then $\neg L \in I$ and $L \in S_P(I)$ which is impossible by definition of self-founded set which requires that $S_P(I) \cup I \in I_P$. So $S_P(I) \cap F_P(I) = \emptyset$ and CW_P is well-defined.

For monotonicity, let $I \subseteq J \in I_P$ and let $L \in \mathrm{CW}_P(I)$. If $L = \neg A$ is a negated atom, then $A \in F_P(I)$ and all clauses with head A contain a body literal which is false in I, hence in J, and we obtain $A \in F_P(J)$. If $L = A$ is an atom, then $A \in S_P(I)$ and there exists a clause $A \leftarrow \texttt{body}$ in $\texttt{ground}(P)$ such that (at least) one of (Si) or (Sii) holds. If (Si) holds, then $\texttt{body}$ is true in I, hence in J, and $A \in S_P(J)$. If (Sii) holds, then some non-negated atoms in $\texttt{body}$ occur in S and all other literals in $\texttt{body}$ are true in I, hence in J, and we obtain $A \in S_P(J)$. ■

The following theorem relates our previous observations to Definition 3.2, in perfect analogy to the correspondence between the stable model semantics, Theorem 2.1, Fages's characterization from Theorem 2.2, the well-founded semantics, and the alternating fixed point characterization. That such analogies are not purely coincidental was already hinted at in the work of Fitting, Denecker, Marek, and Truszczynski [Fit91, Fit02, DMT00].

5.5 Theorem Let P be a normal program and $M_P = \mathsf{lfp}(\mathrm{CW}_P)$. Then the following hold.

[4] Again, it is not really a duality.

(i) M_P is the greatest model among all models I of P such that there is an I-partial level mapping l for P such that P satisfies (CW) with respect to I and l.

(ii) $M_P = \mathsf{lfp}\left(\mathrm{CGL}_P^2\right) \cup \neg\left(B_P \setminus \mathsf{gfp}\left(\mathrm{CGL}_P^2\right)\right)$.

Proof: (i) Let $M_P = \mathsf{lfp}(\mathrm{CW}_P)$ and define the M_P-partial level mapping l_P as follows: $l_P(A) = \alpha$, where α is the least ordinal such that A is not undefined in $\mathrm{CW}_P\!\uparrow\!(\alpha+1)$. The proof will be established by showing the following facts: (1) P satisfies (CW) with respect to M_P and l_P. (2) If I is a model of P and l is an I-partial level mapping such that P satisfies (CW) with respect to I and l, then $I \subseteq M_P$.

(1) Let $A \in \mathsf{dom}(l_P)$ and $l_P(A) = \alpha$. We consider two cases.

(Case i) If $A \in M_P$, then $A \in S_P(\mathrm{CW}_P\!\uparrow\!\alpha)$, hence there exists a clause $A \leftarrow \mathsf{body}$ in $\mathsf{ground}(P)$ such that (Si) or (Sii) holds with respect to $\mathrm{CW}_P\!\uparrow\!\alpha$. If (Si) holds, then all literals in body are true in $\mathrm{CW}_P\!\uparrow\!\alpha$, hence have level less than $l_P(A)$ and (Ci) is satisfied. If (Sii) holds, then some non-negated atoms from body occur in $S_P(\mathrm{CW}_P\!\uparrow\!\alpha)$, hence have level less than or equal to $l_P(A)$, and all remaining literals in body are true in $\mathrm{CW}_P\!\uparrow\!\alpha$, hence have level less than $l_P(A)$. Consequently, A satisfies (Ci) with respect to M_P and l_P.

(Case ii) If $\neg A \in M_P$, then $A \in F_P(\mathrm{CW}_P\!\uparrow\!\alpha)$, hence for all clauses $A \leftarrow \mathsf{body}$ in $\mathsf{ground}(P)$ there exists $L \in \mathsf{body}$ with $\neg L \in \mathrm{CW}_P\!\uparrow\!\alpha$ and $l_P(L) < \alpha$, hence $\neg L \in M_P$. Consequently, A satisfies (Fii) with respect to M_P and l_P, and we have established that fact (1) holds.

(2) We show via transfinite induction on $\alpha = l(A)$, that whenever $A \in I$ (respectively, $\neg A \in I$), then $A \in \mathrm{CW}_P\!\uparrow\!(\alpha+1)$ (respectively, $\neg A \in \mathrm{CW}_P\!\uparrow\!(\alpha+1)$). For the base case, note that if $l(A) = 0$, then $\neg A \in I$ implies that there is no clause with head A in $\mathsf{ground}(P)$, hence $\neg A \in \mathrm{CW}_P\!\uparrow\!1$. If $A \in I$ then consider the set S of all atoms B with $l(B) = 0$ and $B \in I$. We show that S is a self-founded set of P with respect to $\mathrm{CW}_P\!\uparrow\!0 = \emptyset$, and this suffices since it implies $A \in \mathrm{CW}_P\!\uparrow\!1$ by the fact that $A \in S$. So let $C \in S$. Then $C \in I$ and C satisfies condition (Ci) with respect to I and l, and since $l(C) = 0$, we have that there is a definite clause with head C whose body atoms (if it has any) are all of level 0 and contained in I. Hence condition (Sii) (or (Si)) is satisfied for this clause and S is a self-founded set of P with respect to I. So assume now that the induction hypothesis holds for all $B \in B_P$ with $l(B) < \alpha$, and let A be such that $l(A) = \alpha$. We consider two cases.

(Case i) If $A \in I$, consider the set S of all atoms B with $l(B) = \alpha$ and $B \in I$. We show that S is a self-founded set of P with respect to $\mathrm{CW}_P\!\uparrow\!\alpha$, and this suffices since it implies $A \in \mathrm{CW}_P\!\uparrow\!(\alpha+1)$ by the fact that $A \in S$. First note that $S \subseteq I$, so $S \cup I \in I_P$. Now let $C \in S$. Then $C \in I$ and C satisfies condition (Ci) with respect to I and l, so there is a clause $A \leftarrow A_1, \ldots, A_n, \neg B_1, \ldots, \neg B_m$ in $\mathsf{ground}(P)$ such that $A_i, \neg B_j \in I$, $l(A) \geq l(A_i)$, and $l(A) > l(B_j)$ for all i and j. By induction hypothesis we obtain $\neg B_j \in \mathrm{CW}_P\!\uparrow\!\alpha$. If $l(A_i) < l(A)$ for some A_i then we have $A_i \in \mathrm{CW}_P\!\uparrow\!\alpha$, also by induction hypothesis. If there is no A_i with $l(A_i) = l(A)$, then (Si) holds, while $l(A_i) = l(A)$ implies $A_i \in S$, so (Sii) holds.

(Case ii) If $\neg A \in I$, then A satisfies (Fii) with respect to I and l. Hence for all clauses $A \leftarrow \mathsf{body}$ in $\mathsf{ground}(P)$ we have that there is $L \in \mathsf{body}$ with $\neg L \in I$ and $l(L) < \alpha$. Hence for all these L we have $\neg L \in \mathrm{CW}_P\!\uparrow\!\alpha$ by induction hypothesis, and consequently for all clauses $A \leftarrow \mathsf{body}$ in $\mathsf{ground}(P)$ we obtain that body is false in $\mathrm{CW}_P\!\uparrow\!\alpha$ which yields $\neg A \in \mathrm{CW}_P\!\uparrow\!(\alpha+1)$. This establishes fact (2) and concludes the proof of (i).

(ii) We first introduce some notation. Let

$$
\begin{aligned}
L_0 &= \emptyset, \\
G_0 &= B_P, \\
L_{\alpha+1} &= \mathrm{CGL}_P(G_\alpha) \qquad \text{for any ordinal } \alpha, \\
G_{\alpha+1} &= \mathrm{CGL}_P(L_\alpha) \qquad \text{for any ordinal } \alpha, \\
L_\alpha &= \bigcup_{\beta<\alpha} L_\beta \qquad \text{for limit ordinal } \alpha, \\
G_\alpha &= \bigcap_{\beta<\alpha} G_\beta \qquad \text{for limit ordinal } \alpha, \\
L_P &= \mathsf{lfp}(\mathrm{CGL}_P^2), \\
G_P &= \mathsf{gfp}(\mathrm{CGL}_P^2).
\end{aligned}
$$

By transfinite induction, it is easily checked that $L_\alpha \subseteq L_\beta \subseteq G_\beta \subseteq G_\alpha$ whenever $\alpha \leq \beta$. So $L_P = \bigcup L_\alpha$ and $G_P = \bigcap G_\alpha$.

Let $M = L_P \cup \neg(B_P \setminus G_P)$. We intend to apply (i) and first define an M-partial level mapping l. We will take as image set of l, pairs (α, γ) of ordinals, with the lexicographic ordering. This can be done without loss of generality since any set of such pairs, under the lexicographic ordering, is well-ordered, and therefore order-isomorphic to an ordinal. For $A \in L_P$, let $l(A)$ be the pair $(\alpha, 0)$, where α is the least ordinal such that $A \in L_{\alpha+1}$. For $B \notin G_P$, let $l(B)$ be the pair (β, γ), where β is the least ordinal such that $B \notin G_{\beta+1}$, and γ is least such that $B \notin T_{P/L_\beta}\!\downarrow\!\gamma$. It is easily shown that l is well-defined, and we show next by transfinite induction that P satisfies (CW) with respect to M and l.

Let $A \in L_1 = \mathsf{gfp}\left(T^+_{P/B_P}\right)$. Since P/B_P contains exactly all clauses from $\mathsf{ground}(P)$ which contain no negation, we have that A is contained in the greatest two-valued model of a definite subprogram of P, namely P/B_P. So there must be a definite clause in $\mathsf{ground}(P)$ with head A whose corresponding body atoms are also true in L_1, which, by definition of l, must have the same level as A, hence (Ci) is satisfied. Now let $\neg B \in \neg(B_P \setminus G_P)$ such that $B \in (B_P \setminus G_1) = B_P \setminus \mathsf{gfp}\left(T^+_{P/\emptyset}\right)$. Since $P/\emptyset$ contains all clauses from $\mathsf{ground}(P)$ with all negative literals removed, we obtain that B is not contained in the greatest two-valued model of the definite program $P/\emptyset$, and (Fii) is satisfied by Theorem 4.1 using a simple induction argument.

Assume now that, for some ordinal α, we have shown that A satisfies (CW) with respect to M and l for all $A \in B_P$ with $l(A) < (\alpha, 0)$.

Let $A \in L_{\alpha+1} \setminus L_\alpha = \mathsf{gfp}\left(T^+_{P/G_\alpha}\right) \setminus L_\alpha$. Then $A \in \left(T^+_{P/G_\alpha}\!\downarrow\!\gamma\right) \setminus L_\alpha$ for some γ; note that all (negative) literals which were removed by the Gelfond-Lifschitz transformation from clauses with head A have level less than $(\alpha, 0)$. Then A satisfies (Ci) with respect to M and l by definition of l.

Let $A \in (B_P \setminus G_{\alpha+1}) \cap G_\alpha$. Then $A \notin \mathsf{gfp}\left(T^+_{P/L_\alpha}\right)$ and we conclude again from Theorem 4.1, using a simple induction argument, that A satisfies (CW) with respect to M and l.

This finishes the proof that P satisfies (CW) with respect to M and l. It remains to show that M is greatest with this property.

So assume that $M_1 \supset M$ is the greatest model such that P satisfies (CW) with respect to M_1 and some M_1-partial level mapping l_1. Assume $L \in M_1 \setminus M$ and, without loss of generality, let the literal L be chosen such that $l_1(L)$ is minimal. We consider two cases.

(Case i) If $L = \neg A \in M_1 \setminus M$ is a negated atom, then by (Fii) for each clause $A \leftarrow L_1, \dots, L_n$ in $\mathsf{ground}(P)$ there exists i with $\neg L_i \in M_1$ and $l_1(A) > l_1(L_i)$. Hence, $\neg L_i \in M$ and consequently for each clause $A \leftarrow \mathtt{body}$ in P/L_P we have that some atom in $\mathtt{body}$ is false in $M = L_P \cup \neg(B_P \setminus G_P)$. But then $A \notin \mathrm{CGL}_P(L_P) = G_P$, hence $\neg A \in M$, contradicting $\neg A \in M_1 \setminus M$.

(Case ii) If $L = A \in M_1 \setminus M$ is an atom, then $A \notin M = L_P \cup \neg(B_P \setminus G_P)$ and in particular $A \notin L_P = \mathsf{gfp}\left(T^+_{P/G_P}\right)$. Hence $A \notin T^+_{P/G_P} \downarrow \gamma$ for some γ, which can be chosen to be least with this property. We show by induction on γ that this leads to a contradiction, to finish the proof.

If $\gamma = 1$, then there is no clause with head A in P/G_P, i.e. for all clauses $A \leftarrow \mathtt{body}$ in $\mathsf{ground}(P)$ we have that $\mathtt{body}$ is false in M, hence in M_1, which contradicts $A \in M_1$.

Now assume that there is no $B \in M_1 \setminus M$ with $B \notin T^+_{P/G_P} \downarrow \delta$ for any $\delta < \gamma$, and let $A \in M_1 \setminus M$ with $A \notin T^+_{P/G_P} \downarrow \gamma$, which implies that γ is a successor ordinal. By $A \in M_1$ and (Ci) there must be a clause $A \leftarrow A_1, \dots, A_n \neg B_1, \dots, \neg B_m$ in $\mathsf{ground}(P)$ with $A_i, \neg B_j \in M_1$ for all i and j. However, since $A \notin T^+_{P/G_P} \downarrow \gamma$ we obtain that for each $A \leftarrow A_1, \dots, A_n$ in P/G_P, hence for each $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in $\mathsf{ground}(P)$ with $\neg B_1, \dots, \neg B_m \in \neg(B_P \setminus G_P) \subseteq M \subseteq M_1$ there is A_i with $A_i \notin T^+_{P/G_P} \downarrow (\gamma - 1) \subseteq M$, and by induction hypothesis we obtain $A_i \notin M_1$. So $A_i \in M_1$ and $A_i \notin M_1$, which is a contradiction and concludes the proof. ∎

5.6 Definition For a normal program P, we call $\mathsf{lfp}(\mathrm{CW}_P)$ the *maximally circular well-founded model* (*maxwf model*) of P.

6 Circular Beliefs

Observe the program P consisting of the single clause $p \leftarrow p$. In classical logic, this clause is equivalent to $p \vee \neg p$, which has two models, namely $\{p\}$ and $\{\neg p\}$. We note that $\{\neg p\}$ coincides with the stable model of P, while $\{p\}$ is its maxstable model. So we refer to stable models as *minimally circular stable models* (*minstable models*). For the program from Example 3.1, we have that M_2^+ is the unique minstable model, while M_1^+ is the unique maxstable model.

6.1 Example Consider the following program, where X is a variable, and `bob` a constant.

$$\begin{aligned} \mathtt{penguin}(X) &\leftarrow \mathtt{penguin}(X), \mathtt{bird}(X) \\ \mathtt{flies}(X) &\leftarrow \mathtt{bird}(X), \neg\mathtt{penguin}(X) \\ \mathtt{bird}(\mathtt{bob}) &\leftarrow \end{aligned}$$

The first clause is supposed to represent the knowledge that a bird *may* be a penguin. The only minstable model of the program is $\{\mathtt{bird}(\mathtt{bob}), \mathtt{flies}(\mathtt{bob})\}$, i.e. `bob` does fly. The only maxstable model of the program is $\{\mathtt{bird}(\mathtt{bob}), \mathtt{penguin}(\mathtt{bob})\}$, i.e. `bob` does not fly, but is a penguin. Both the maxstable and the minstable model describe possible settings if the first clause is interpreted as indicated.

Consider the program consisting of the two clauses $p \leftarrow p$ and $q \leftarrow q$, and we would like to interpret both clauses as before, namely as allowing a choice for believing p, or q, or both, or none. Thus we have four possibilities, which are not all captured by the minstable model $\emptyset$ and the maxstable model $\{p, q\}$. We rectify this situation by introducing the following definitions, inspired by our investigations in the previous chapters.

6.2 Definition Let P be a normal logic program, I be a model of P, and l be an I-partial level mapping for P. We say that P *satisfies* (C) *with respect to* I *and* l, if each $A \in \mathsf{dom}(l)$ satisfies one of the following conditions.

(Ci) $A \in I$ and there exists a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in $\mathsf{ground}(P)$ with $A_i, \neg B_j \in I$, $l(A) \geq l(A_i)$, and $l(A) > l(B_j)$ for all i and j.

(Cii) $\neg A \in I$ and for each clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in $\mathsf{ground}(P)$ (at least) one of the following conditions holds:

- (Ciia) There exists $i \in \{1, \dots, n\}$ with $\neg A_i \in I$ and $l(A) \geq l(A_i)$.
- (Ciib) There exists $j \in \{1, \dots, m\}$ with $B_j \in I$ and $l(A) > l(B_j)$.

If $A \in \mathsf{dom}(l)$ satisfies (Ci), then we say that A *satisfies* (Ci) *with respect to* I *and* l, and similarly if $A \in \mathsf{dom}(l)$ satisfies (Cii). If M is a model of P which is maximal under the condition that P satisfies (C) with respect to M and some M-partial level mapping l, then we call M a *circular model* of P. The corresponding semantics is called the *cirular semantics* or[5] the *switch semantics* of P.

For the program consisting of the two clauses $p \leftarrow p$ and $q \leftarrow q$, we have four (total) circular models $\{\neg p, \neg q\}$, $\{p, \neg q\}$, $\{\neg p, q\}$, and $\{p, q\}$, as the reader will easily verify. Again, in this simple case the models coincide with the models of $(p \leftarrow p) \wedge (q \leftarrow q)$ in classical logic. Using the stable model semantics for extended disjunctive logic programs as introduced by Gelfond and Lifschitz [GL91], the same could be achieved using the program consisting of the two clauses $p \vee \neg p \leftarrow$ and $q \vee \neg q \leftarrow$. The following example is less abstract, but conveys the same idea.

6.3 Example Consider a coffee delivery robot, which has two choices of doors for entering the room where the coffee is to be delivered. It is unknown to the robot whether these doors, or one of them, are open.

$$\begin{aligned}
\mathtt{open}(X) &\leftarrow \mathtt{open}(X), \mathtt{door}(X)\\
\mathtt{deliverable} &\leftarrow \mathtt{open}(X)\\
\mathtt{undeliverable} &\leftarrow \neg\mathtt{deliverable}\\
\mathtt{door}(1) &\leftarrow\\
\mathtt{door}(2) &\leftarrow
\end{aligned}$$

[5]The latter term is a proposal by Matthias Wendt.

This program has four (total) circular models, as follows.

$$\{\mathtt{door}(1), \mathtt{door}(2), \mathtt{open}(1), \mathtt{open}(2), \mathtt{deliverable}\} \cup \neg\{\mathtt{undeliverable}\}$$

(The maxstable model.)

$$\{\mathtt{door}(1), \mathtt{door}(2), \mathtt{undeliverable}\} \cup \neg\{\mathtt{open}(1), \mathtt{open}(2), \mathtt{deliverable}\}$$

(The minstable model.)

$$\{\mathtt{door}(1), \mathtt{door}(2), \mathtt{open}(1), \mathtt{deliverable}\} \cup \neg\{\mathtt{open}(2), \mathtt{undeliverable}\}$$

$$\{\mathtt{door}(1), \mathtt{door}(2), \mathtt{open}(2), \mathtt{deliverable}\} \cup \neg\{\mathtt{open}(1), \mathtt{undeliverable}\}$$

It is remarkable that this result can be achieved without the use of disjunctive logic programming, as e.g. in [GL91].

Let us shortly discuss our examples. Under the circular semantics, a program in general has several distinguished three-valued models, which we call circular models. In essence, these different models come from interpreting the *switch* $p \leftarrow p$ similarly to classical logic, i.e. as classical disjunction. However, this dependency of an atom on itself may be hidden in the program, e.g. within two clauses $a \leftarrow b$ and $b \leftarrow a$. In this latter case the only two circular models are $\{a, b\}$ and $\{\neg a, \neg b\}$.

As a result we obtain a setting in which we can model choices or uncertain knowledge about the environment — whether a door is open or not — *without* the use of explicit disjunction. It is certainly to be investigated whether the circular semantics is a reasonable framework for knowledge representation and reasoning purposes. However, our main concern in this paper was to show how the methodology from [HW02b, HW02a] can be used to create new semantics, which is are least potentially meaningful. So we restrict our analysis of circular models to the examples above and the following theorems, which shed some additional light on circular models.

6.4 Theorem Let P be a normal program and N be a total circular model of P. Then $M = N^+$ is a fixed point of $T^+_{P/M}$. The converse does not hold in general.

Proof: We have to show that $M = T^+_{P/M}(M)$. So let $A \in M$. Then by (Ci) there exists a clause $A \leftarrow A_1, \dots, A_n, \neg B_1, \dots, \neg B_m$ in $\mathsf{ground}(P)$ with $A_i \in M$ and $B_J \notin M$, for all i and j. Thus, there exists a clause $A \leftarrow A_1, \dots, A_n$ in P/M with $A_i \in M$ for all i, and we conclude that $A \in T^+_{P/M}(M)$, and we have shown that $M \subseteq T^+_{P/M}(M)$. Now let $A \notin M$. Then by (Cii) we have that for each clause $A \leftarrow \mathsf{body}$ in $\mathsf{ground}(P)$ (at least) one of (Ciia) or (Ciib) holds with respect to M. In either case, we have that body is false in M, and we conclude $A \notin T^+_{P/M}(M)$, showing $M \supseteq T^+_{P/M}(M)$.

In order to see that the converse does not hold in general, consider the program consisting of the two clauses $p \leftarrow \neg q$ and $q \leftarrow \neg p$, which has the single circular model $\emptyset$, but $T^+_{P/\{p\}}(\{p\}) = \{p\}$. ■

The Fitting model of a program approximates all circular models of a program, as follows.

6.5 Theorem Let P be a program, F be its Fitting model, and M be a circular model of P. Then $F \subseteq M$. Furthermore, there exists a program P, the Fitting model F of which is not circular, but F is the intersection of all circular models of P.

Proof: It follows immediately from Definitions 2.3 and 6.2 that P satisfies (C) with respect to F and l, which suffices.

For the second assertion consider the program P consisting of the single clause $p \leftarrow p$. Then $\emptyset$ is the Fitting model of P, while its circular models are $\{p\}$ and $\{\neg p\}$, which shows the claim. ■

Finally, we note that the well-founded model of a program is not necessarily circular. Indeed, consider the program P consisting of the two clauses $p \leftarrow p$ and $p \leftarrow \neg p$. Then $\emptyset$ is the well-founded model of P while $\{p\}$ is its maxwf model. So the well-founded model is not maximal as a model satisfying (C). Likewise, the maxwf model is not necessarily circular, which can be seen from the program Q consisting of the single clause $q \leftarrow q, \neg q$, which has well-founded model $\{\neg q\}$ and maxwf model $\emptyset$. Note also that the program $P \cup Q$ has well-founded model $\{\neg q\}$ and maxwf model $\{p\}$, while its unique circular model is $\{p, \neg q\}$. It is trivial, however, to verify that in general both the well-founded and the maxwf model satisfy (C) with respect to the given program, although maximality is not guaranteed as the examples just given show.

7 Conclusions and Further Work

We have displayed the usefulness of the approach from [HW02b, HW02a] for creating new semantics, which are potentially meaningful. Using solely formal arguments we have designed a semantics which allows to deal with circular belief and uncertainty, in a novel way. Circular reasoning, although judged undesirable in scientific discourse, is a natural ingredient in the way humans think and argue in everyday life, and we refer to [Rip02] for a recent publication on this issue. Certainly, it remains to be determined whether the proposed circular semantics is useful for knowledge representation and reasoning purposes. Our main motivation, however, was to study the methodology proposed in [HW02b, HW02a] by providing a case in point which shows how it can lead to the creation of new semantics. Indeed, we note that *all* proofs of theorems in Sections 4 and 5 follow the general proof scheme laid out in [HW02a]. The work presented here is explorative, but we expect that our studies will eventually lead to a meta-theory based on [HW02b, HW02a].

References

[ABW88] Krzysztof R. Apt, Howard A. Blair, and Adrian Walker. Towards a theory of declarative knowledge. In Jack Minker, editor, *Foundations of Deductive Databases and Logic Programming*, pages 89–148. Morgan Kaufmann, Los Altos, CA, 1988.

[AP93] Krzysztof R. Apt and Dino Pedreschi. Reasoning about termination of pure prolog programs. *Information and Computation*, 106:109–157, 1993.

[DMT00] Marc Denecker, V. Wiktor Marek, and Miroslaw Truszczynski. Approximating operators, stable operators, well-founded fixpoints and applications in non-monotonic reasoning. In Jack Minker, editor, *Logic-based Artificial Intelligence*, chapter 6, pages 127–144. Kluwer Academic Publishers, Boston, 2000.

[Fag94] François Fages. Consistency of Clark's completion and existence of stable models. *Journal of Methods of Logic in Computer Science*, 1:51–60, 1994.

[Fit85] Melvin Fitting. A Kripke-Kleene-semantics for general logic programs. *The Journal of Logic Programming*, 2:295–312, 1985.

[Fit91] Melvin Fitting. Bilattices and the semantics of logic programming. *The Journal of Logic Programming*, 11:91–116, 1991.

[Fit02] Melvin Fitting. Fixpoint semantics for logic programming — A survey. *Theoretical Computer Science*, 278(1–2):25–51, 2002.

[GL88] Michael Gelfond and Vladimir Lifschitz. The stable model semantics for logic programming. In Robert A. Kowalski and Kenneth A. Bowen, editors, *Logic Programming. Proceedings of the 5th International Conference and Symposium on Logic Programming*, pages 1070–1080. MIT Press, 1988.

[GL91] Michael Gelfond and Vladimir Lifschitz. Classical negation in logic programs and disjunctive databases. *New Generation Computing*, 9:365–385, 1991.

[HW02a] Pascal Hitzler and Matthias Wendt. A uniform approach to logic rogramming semantics. Technical Report WV–02–14, Knowledge Representation and Reasoning Group, Artificial Intelligence Institute, Department of Computer Science, Dresden University of Technology, Dresden, Germany, 2002. Submitted.

[HW02b] Pascal Hitzler and Matthias Wendt. The well-founded semantics is a stratified Fitting semantics. In Matthias Jarke, Jana Koehler, and Gerhard Lakemeyer, editors, *Proceedings of the 25th Annual German Conference on Artificial Intelligence, KI2002, Aachen, Germany, September 2002*, volume 2479 of *Lecture Notes in Artificial Intelligence*, pages 205–221. Springer, Berlin, 2002.

[Jac01] Jacek Jachymski. Order-theoretic aspects of metric fixed-point theory. In William A. Kirk and Brailey Sims, editors, *Handbook of Metric Fixed Point Theory*, pages 613–641. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2001.

[Kun87] Kenneth Kunen. Negation in logic programming. *The Journal of Logic Programming*, 4:289–308, 1987.

[Lif02] Vladimir Lifschitz. Answer set programming and plan generation. *Artificial Intelligence*, 138:39–54, 2002.

[Llo88] John W. Lloyd. *Foundations of Logic Programming.* Springer, Berlin, 1988.

[MT99] V. Wiktor Marek and Miroslav Truszczyński. Stable models and an alternative logic programming paradigm. In Krzysztof R. Apt, V. Wiktor Marek, Miroslav Truszczyński, and David S. Warren, editors, *The Logic Programming Paradigm: A 25-Year Persepective*, pages 375–398. Springer, Berlin, 1999.

[PP90] Halina Przymusinska and Teodor C. Przymusinski. Weakly stratified logic programs. *Fundamenta Informaticae*, 13:51–65, 1990.

[Prz88] Teodor C. Przymusinski. On the declarative semantics of deductive databases and logic programs. In Jack Minker, editor, *Foundations of Deductive Databases and Logic Programming*, pages 193–216. Morgan Kaufmann, Los Altos, CA, 1988.

[Rip02] Lance J. Rips. Circular reasoning. *Cognitive Science*, 26:767–795, 2002.

[SHLG94] Viggo Stoltenberg-Hansen, Ingrid Lindström, and Edward R. Griffor. *Mathematical Theory of Domains*. Cambridge University Press, 1994.

[SNS0x] Patrik Simons, Ilkka Niemelä, and Timo Soininen. Extending and implementing the stable model semantics. *Artificial Intelligence*, 200x. To appear.

[vG89] Allen van Gelder. The alternating fixpoint of logic programs with negation. In *Proceedings of the Eighth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, Philadelphia, Pennsylvania*, pages 1–10. ACM Press, 1989.

[vGRS91] Allen van Gelder, Kenneth A. Ross, and John S. Schlipf. The well-founded semantics for general logic programs. *Journal of the ACM*, 38(3):620–650, 1991.