Autoepistemic Answer Set Programming

Pedro Cabalar[a], Jorge Fandinno[b], Luis Fariñas del Cerro[c]

[a] *University of Corunna, SPAIN*
`cabalar@udc.es`
[b] *University of Potsdam, GERMANY*
`fandinno@uni-potsdam.de`
[c] *Institut de Recherche en Informatique de Toulouse*
Toulouse Universtiy, CNRS, FRANCE
`farinas@irit.fr`

Abstract

Defined by Gelfond in 1991-94 (G94), epistemic specifications constitute an extension of Answer Set Programming (ASP) that introduces *subjective literals*. A subjective literal allows checking whether some regular literal is true in all (or in some of) the answer sets of the program, that are further collected in a set called *world view*. One epistemic program may yield several world views but, under the original G94 semantics, some of them resulted from self-supported derivations. During the last eight years, several alternative approaches have been proposed to get rid of these self-supported world views. Unfortunately, their success could only be measured by studying their behaviour on a set of common examples in the literature, since no formal property of "self-supportedness" had been defined. To fill this gap, we extend in this paper the idea of unfounded set from standard logic programming to the epistemic case. We define when a world view is *founded* with respect to some program. Accordingly, we define the *foundedness* property for an arbitrary semantics, so it holds when its world views are always founded. Using counterexamples, we explain that the previous approaches violate foundedness, and proceed to propose a new semantics based on a combination of Moore's Autoepistemic Logic and Pearce's Equilibrium Logic. This combination paves the way for the development of an autoepistemic extension of ASP. The main result proves that this new semantics precisely captures the set of founded G94 world views.

Keywords: Answer Set Programming; Epistemic Specifications; Epistemic Logic Programs; Autoepistemic Logic; Non-Monotonic Reasoning;

1 Introduction

Epistemic reasoning [32, 18] constitutes a crucial feature for any agent to be considered intelligent. The capacity of representing and reasoning about knowledge and beliefs has proved to be a key property in different domains such as planning under incomplete information, speech acts in natural language understanding, software verification of security protocols, formalisation of multi-agent systems or foundations of game theory (see [13, 17, 3, 15, 52, 60]). An important field in Knowledge Representation (KR) where epistemic reasoning has played a relevant role since its inception is *non-monotonic reasoning*. There, default rules have been frequently addressed in terms of modal constructions expressing the agent's own knowledge and beliefs, as a kind of introspection. There exists a vast literature on non-monotonic modal logics (see for instance [44, 40, 43, 38]) among which Moore's *Autoepistemic Logic* (AEL) [47] is perhaps the most prominent and well-studied approach for non-monotonic epistemic introspection. Moreover, AEL has been commonly used in translations or encodings for other non-monotonic approaches, like default negation in logic programming [28, 39].

Despite of its clear significance in KR, the impact of epistemic reasoning in practical applications has been moderate so far. One possible reason is that dealing with the agent's knowledge normally implies an increase in computational complexity. However, a more important obstacle appears when we extend an existing KR formalism with epistemic constructs and there exist multiple options for their interpretation without a clear orientation or agreement about which one *preserves the main features of the extended formalism* in the best way. This is precisely the situation in the case of *Answer Set Programming* (ASP) [42, 49], one of the most popular paradigms for practical KR and problem solving based on the *stable model* [25] semantics for disjunctive logic programs.

The first steps towards an epistemic extension of answer set programming can be traced back to the language of *epistemic specifications*. This language was proposed by Gelfond in three consecutive papers [23, 26, 29] and extends ASP with epistemic operators **K** and **M**. Using these constructs, it is possible to check whether a regular literal l is true in *every* stable model (written $\mathbf{K}\, l$)

or in *some* stable model (written $\mathbf{M}\, l$) of the program. For instance, the rule:

$$a \leftarrow \neg \mathbf{K}\, b \tag{1}$$

means that a must hold if we cannot prove that all the stable models contain b. The definition of a "satisfactory" semantics for epistemic specifications has proved to be a non-trivial enterprise, as shown by the list of different attempts proposed so far [21, 23, 24, 34, 56, 58, 59, 62]. The main difficulty arises because subjective literals query the set of stable models but, at the same time, occur in rules that determine those stable models. As an example, the program consisting of:

$$b \leftarrow \neg \mathbf{K}\, a \tag{2}$$

and (1) has now two rules defining atoms a and b in terms of the presence of those same atoms in all the stable models. To solve this kind of cyclic interdependence, the original semantics by Gelfond [23, 27, 29] (abbreviated[1] as G94) considered different alternative *world views* or sets of stable models. In the case of program (1)-(2), G94 yields two alternative world views[2], $[\{a\}]$ and $[\{b\}]$, each one containing a single stable model, and this is also the behaviour obtained in the remaining approaches developed later on. The feature that made G94 unconvincing, though, was the generation of self-supported world views. A prototypical example for this effect is the epistemic program consisting of the single rule:

$$a \leftarrow \mathbf{K}\, a \tag{3}$$

whose world views under G94 are $[\emptyset]$ and $[\{a\}]$. The latter is considered counter-intuitive by all authors[3] because it relies on a self-supported derivation: a is derived from $\mathbf{K}\, a$ by rule (3), but the only way to obtain $\mathbf{K}\, a$ is rule (3) itself. Although the rejection of world views of this kind seems natural, the truth is that all approaches in the literature have concentrated on

[1]As a notation convention, we abbreviate each semantics name using its original publication, with the initial of the first author's last name followed by the last two digits of the publication year.

[2]For the sake of readability, sets of propositional interpretations are embraced with [ ] instead of { }.

[3]This includes Gelfond himself, who used this same example to propose a new variant in [24], further modified in [34] later on.

studying the effects on individual examples, rather than capturing the absence of self-supportedness as a formal property. To achieve such a goal, we would need to establish some kind of *derivability* condition in a very similar fashion as done with *unfounded sets* [61] for standard logic programs. To understand the similarity, think about the (tautological) rule $a \leftarrow a$. The classical models of this rule are $\emptyset$ and $\{a\}$, but the latter cannot be a stable model because a is not derivable applying the rule. Intuitively, an unfounded set is a collection of atoms that is not derivable from a given program and a fixed set of assumptions, as happens to $\{a\}$ in the last example. As proved by Leone et al. [36], the stable models of any disjunctive logic program are precisely its classical models that are *founded*, that is, that do not admit any unfounded set. As we can see, the situation in (3) is pretty similar to $a \leftarrow a$ but, this time, involves derivability through subjective literals. An immediate option is, therefore, extending the definition of unfounded sets for the case of epistemic programs – this constitutes, indeed, the first contribution of this paper.

Once the property of *founded* world views is explicitly stated, the paper proposes a new semantics for epistemic specifications, called *Founded Autoepistemic Equilibrium Logic* (FAEEL), that satisfies that property. In the spirit of [21, 58, 62], our proposal actually constitutes a full modal non-monotonic logic where $\mathbf{K}$ becomes the usual necessity operator applicable to arbitrary formulas. Formally, FAEEL is a combination of Pearce's *Equilibrium Logic* [50, 51], a well-known logical characterisation of stable models, with Moore's AEL, one of the most representative approaches among modal non-monotonic logics. The reason for choosing Equilibrium Logic is quite obvious, as it has proved its utility for characterising other extensions of ASP [1, 2, 4, 5, 7, 8, 9, 10, 12, 14, 20, 30, 53], including the already mentioned epistemic approaches [21, 58, 62]. As for the choice of AEL, it shares with epistemic specifications the common idea of *agent's introspection* where $\mathbf{K}\,\varphi$ means that φ is one of the agent's beliefs. The only difference is that those beliefs are just classical models in the case of AEL whereas epistemic specifications deal with stable models instead. Interestingly, the problem of self-supported models has also been extensively studied in AEL [35, 41, 48, 54], where the formula $\mathbf{K}\,a \rightarrow a$, analogous to (3), also yields an unfounded world view[4] $[\{a\}]$. Our solution consists in combining the monotonic bases of AEL

[4]Technically, AEL is defined in terms of *theory expansions* but each one can be char-

and Equilibrium Logic (the modal logic KD45 and the intermediate logic of Here-and-There (HT) [31], respectively), but defining a two-step models selection criterion that simultaneously keeps the agent's beliefs as stable models and avoids unfounded world views from the use of the modal operator **K**. As expected, we prove that FAEEL guarantees the property of founded world views, among other properties lifted from standard ASP. Our main result, however, goes further and asserts that the FAEEL world views of an epistemic program are precisely the set of founded G94 world views. We reach, in this way, an analogous situation to the case of standard logic programming, where stable models are the set of founded classical models of the program. These results suggest that FAEEL is a solid formal basis for the development of an autoepistemic extension of ASP.

The rest of the paper is organised as follows. Sections 2 and 3 respectively revisit the background knowledge about equilibrium logic and epistemic specifications necessary for the rest of the paper. Section 4 introduces the foundedness property for epistemic logic programs and then, Section 5 provides a pair of counterexamples that suffice to prove that it does not hold for any of the previously existing semantics. In Section 6 we introduce FAEEL and explain some of its properties, making special emphasis on its relation to G94, showing that the latter can be captured as a special subset of FAEEL models. Section 7, contains the proof of the main result, that is, FAEEL-world views are precisely the founded G94-world views. The proof relies on an alternative characterisation of FAEEL that starts from G94 semantics and imposes an additional semantic condition which can be considered as a semantic counterpart of foundedness. In Section 8 we make a comparison among the different semantics, using several examples from the literature and including a table where, apart from foundedness, we also consider other four formal properties recently proposed, showing that only FAEEL satisfies all of them so far. Finally, Section 9 concludes the paper.

2 Background

We begin recalling the basic definitions of equilibrium logic and its relation to stable models. We start from the syntax of propositional logic, with formulas built from combinations of atoms in a set At with operators $\wedge, \vee, \bot$ and $\rightarrow$ in the usual way. We define the derived operators

acterised by a set of classical models with the same form of a world view [46, 55].

$\varphi \leftrightarrow \psi \stackrel{\text{def}}{=} (\varphi \to \psi) \wedge (\psi \to \varphi)$, $(\varphi \leftarrow \psi) \stackrel{\text{def}}{=} (\psi \to \varphi)$, $\neg\varphi \stackrel{\text{def}}{=} (\varphi \to \bot)$ and $\top \stackrel{\text{def}}{=} \neg\bot$.

A *propositional interpretation* T is a set of atoms $T \subseteq At$. We write $T \models \varphi$ to represent that T classically satisfies formula φ. An *HT-interpretation* is a pair $\langle H, T\rangle$ (respectively called "here" and "there") of propositional interpretations such that $H \subseteq T \subseteq At$; it is said to be *total* when $H = T$. We write $\langle H, T\rangle \models \varphi$ to represent that $\langle H, T\rangle$ *satisfies* a formula φ under the recursive conditions:

- $\langle H, T\rangle \not\models \bot$
- $\langle H, T\rangle \models p$ iff $p \in H$
- $\langle H, T\rangle \models \varphi \wedge \psi$ iff $\langle H, T\rangle \models \varphi$ and $\langle H, T\rangle \models \psi$
- $\langle H, T\rangle \models \varphi \vee \psi$ iff $\langle H, T\rangle \models \varphi$ or $\langle H, T\rangle \models \psi$
- $\langle H, T\rangle \models \varphi \to \psi$ iff both:
  (i) $T \models \varphi \to \psi$ and
  (ii) $\langle H, T\rangle \not\models \varphi$ or $\langle H, T\rangle \models \psi$

As usual, we say that $\langle H, T\rangle$ is a *model* of a theory Γ, in symbols $\langle H, T\rangle \models \Gamma$, iff $\langle H, T\rangle \models \varphi$ for all $\varphi \in \Gamma$. It is easy to see that $\langle T, T\rangle \models \Gamma$ iff $T \models \Gamma$ classically. For this reason, we will identify $\langle T, T\rangle$ simply as T and will use '$\models$' equally. By $\mathrm{CL}[\Gamma]$ we denote the set of all classical models of Γ. Interpretation $\langle T, T\rangle = T$ is a *stable (or equilibrium) model* of a theory Γ iff $T \models \Gamma$ and there is no $H \subset T$ such that $\langle H, T\rangle \models \Gamma$. We write $\mathrm{SM}[\Gamma]$ to stand for the set of all stable models of Γ. Note that $\mathrm{SM}[\Gamma] \subseteq \mathrm{CL}[\Gamma]$ by definition.

3 G94 semantics for epistemic theories

In this section we provide a straightforward generalisation of G94 allowing its application to arbitrary modal theories. Formulas are extended with the necessity operator $\mathbf{K}$ according to the following grammar:

$$\varphi ::= \bot \mid a \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \to \varphi_2 \mid \mathbf{K}\,\varphi \qquad \text{for any atom } a \in At.$$

An *(epistemic) theory* is a set of formulas. In our context, the epistemic reading of $\mathbf{K}\,\psi$ is that "ψ is one of the agent's beliefs." Thus, a formula φ is said to be *subjective* if all its atom occurrences (having at least one) are in the scope of $\mathbf{K}$. Analogously, φ is said to be *objective* if $\mathbf{K}$ does not occur in φ. For instance, $\neg\mathbf{K}\,a \vee \mathbf{K}\,b$ is subjective, $\neg a \vee b$ is objective and $\neg a \vee \mathbf{K}\,b$ none of the two.

To represent the agent's beliefs we will use a set $\mathbb{W}$ of propositional interpretations, called *belief view*. Each interpretation $I \in \mathbb{W}$ is further said to be a *belief set*. The difference between belief and knowledge is that the former may not hold in the real world. Thus, satisfaction of formulas will be defined with respect to an interpretation $I \subseteq At$, possibly $I \notin \mathbb{W}$, that accounts for the real world: the pair $(\mathbb{W}, I)$ is called *belief interpretation* (or interpretation in modal logic KD45). Modal satisfaction is also written $(\mathbb{W}, I) \models \varphi$ (ambiguity is removed by the interpretation on the left) and follows the conditions:

- $(\mathbb{W}, I) \not\models \bot$,
- $(\mathbb{W}, I) \models a$ iff $a \in I$, for any atom $a \in At$,
- $(\mathbb{W}, I) \models \psi_1 \wedge \psi_2$ iff $(\mathbb{W}, I) \models \psi_1$ and $(\mathbb{W}, I) \models \psi_2$,
- $(\mathbb{W}, I) \models \psi_1 \vee \psi_2$ iff $(\mathbb{W}, I) \models \psi_1$ or $(\mathbb{W}, I) \models \psi_2$,
- $(\mathbb{W}, I) \models \psi_1 \to \psi_2$ iff $(\mathbb{W}, I) \not\models \psi_1$ or $(\mathbb{W}, I) \models \psi_2$, and
- $(\mathbb{W}, I) \models \mathbf{K}\,\psi$ iff $(\mathbb{W}, J) \models \psi$ for all $J \in \mathbb{W}$.

Notice that implication here is classical, that is, $\varphi \to \psi$ is equivalent to $\neg\varphi \vee \psi$ in this context. A belief interpretation $(\mathbb{W}, I)$ is a *belief model* of Γ iff $(\mathbb{W}, J) \models \varphi$ for all $\varphi \in \Gamma$ and all $J \in \mathbb{W} \cup \{I\}$. We say that $\mathbb{W}$ is an *epistemic model* of Γ, and abbreviate this as $\mathbb{W} \models \Gamma$, iff $(\mathbb{W}, J) \models \varphi$ for all $\varphi \in \Gamma$ and all $J \in \mathbb{W}$. Belief models defined in this way correspond to modal logic KD45 whereas epistemic models correspond to S5.

Example 1. *Take the theory $\Gamma_1 = \{\neg\mathbf{K}\,b \to a\}$ corresponding to rule (1). An epistemic model $\mathbb{W} \models \Gamma_1$ must satisfy: $\langle\mathbb{W}, J\rangle \models \mathbf{K}\,b$ or $\langle\mathbb{W}, J\rangle \models a$, for all $J \in \mathbb{W}$. We get three epistemic models from $\mathbf{K}\,b$, namely, $[\{b\}]$, $[\{a, b\}]$, and $[\{b\}, \{a, b\}]$ and the rest of cases must force atom a to be true, so we also get $[\{a\}]$ and $[\{a\}, \{a, b\}]$. In other words, Γ_1 has the same epistemic models as $\mathbf{K}\,b \vee \mathbf{K}\,a$.* $\square$

Note that rule (1) alone did not seem to provide any reason for believing b, but we got three epistemic models above satisfying $\mathbf{K}\,b$. Thus, we will be only interested in some epistemic models (we will call *world views*) that minimize the agent's beliefs in some sense. To define such a minimisation we rely on the following syntactic transformation provided by Truszczyński [59].

Definition 1 (Subjective reduct)**.** *The* subjective reduct *of a theory* Γ *with respect to a belief view* $\mathbb{W}$*, also written* $\Gamma^{\mathbb{W}}$*, is obtained by replacing each maximal subformula of the form* $\mathbf{K}\,\varphi$*: by* $\top$*, if* $\mathbb{W} \models \mathbf{K}\,\varphi$*; or by* $\bot$*, otherwise. Notice that* $\Gamma^{\mathbb{W}}$ *is a classical, non-modal theory.* □

Finally, we impose a fixpoint condition where, depending on whether each belief set $I \in \mathbb{W}$ is required to be a stable model of the reduct or just a classical model, we get G94 or AEL semantics, respectively.

Definition 2 (AEL and G94 world views)**.** *A belief view* $\mathbb{W}$ *is called an* AEL-world view *of a theory* Γ *iff* $\mathbb{W} = \mathrm{CL}[\Gamma^{\mathbb{W}}]$*, and is called a* G94-world view *of* Γ *iff* $\mathbb{W} = \mathrm{SM}[\Gamma^{\mathbb{W}}]$*.* □

Example 2 (Example 1 revisited)**.** *Take any* $\mathbb{W}$ *such that* $\mathbb{W} \models \mathbf{K}\,b$*. Then,* $\Gamma_1^{\mathbb{W}} = \{\bot \to a\}$ *with* $\mathrm{CL}[\Gamma_1^{\mathbb{W}}] = [\emptyset, \{a\}, \{b\}, \{a,b\}]$ *and* $\mathrm{SM}[\Gamma_1^{\mathbb{W}}] = [\emptyset]$*. None of the two satisfies* $\mathbf{K}\,b$ *so* $\mathbb{W}$ *cannot be fixpoint for G94 or AEL. If* $\mathbb{W} \not\models \mathbf{K}\,b$ *instead, we get* $\Gamma_1^{\mathbb{W}} = \{\top \to a\}$*, whose classical models are* $\{a\}$ *and* $\{a,b\}$*, but only the former is stable. As a result,* $\mathbb{W} = [\{a\}, \{a,b\}]$ *is the unique AEL world view and* $\mathbb{W} = [\{a\}]$ *the unique G94 world view.* □

Example 3. *Take now the theory* $\Gamma_3 = \{\mathbf{K}\,a \to a\}$ *corresponding to rule* (3)*. If* $\mathbb{W} \models \mathbf{K}\,a$ *we get* $\Gamma_3^{\mathbb{W}} = \{\top \to a\}$ *and* $\mathrm{CL}[\Gamma_3^{\mathbb{W}}] = \mathrm{SM}[\Gamma_3^{\mathbb{W}}] = \{a\}$ *so* $\mathbb{W} = [\{a\}]$ *is an AEL and G94 world view. If* $\mathbb{W} \not\models \mathbf{K}\,a$*, the reduct becomes* $\Gamma_3^{\mathbb{W}} = \{\bot \to a\}$*, a classical tautology with unique stable model* $\emptyset$*. As a result,* $\mathbb{W} = [\emptyset, \{a\}]$ *is the other AEL world view, while* $\mathbb{W} = [\emptyset]$ *is the second G94 world view.* □

As we can see, the difference between AEL and G94 is that we respectively use classical $\mathrm{CL}[\Gamma^{\mathbb{W}}]$ versus stable $\mathrm{SM}[\Gamma^{\mathbb{W}}]$ models of the reduct $\Gamma^{\mathbb{W}}$. It is well known that adding the excluded middle axiom:

$$a \vee \neg a \tag{4}$$

for all atoms $a \in At$ makes equilibrium logic collapse into classical logic. This leads us to the next result.

Proposition 1. $\mathbb{W}$ *is an AEL world view of some theory* Γ *iff* $\mathbb{W}$ *is a G94-world view of* $\Gamma \cup (4)$*.* □

To the best of our knowledge, this connection was not presented so far in the literature. As a consequence of this result, any solver for epistemic specifications under G94 semantics can be easily adapted to compute AEL world views: it suffices with adding the excluded middle axiom (4) for each atom in the signature. In fact, this axiom is easily implemented in modern ASP solvers by resorting to a so-called *choice rule* of the form "$0\{a\}1$" or simply "$\{a\}$." Moreover, in this way, we may even decide to select some part of the signature to behave as AEL (adding the corresponding choice rules) and the rest of atoms to follow the standard closed world assumption in ASP, more convenient for expressing defaults or inductive relations.

4 Founded world views of epistemic specifications

As we explained in the introduction, world view $[\{a\}]$ of $\{\mathbf{K}\, a \rightarrow a\}$ is considered to be "self-supported" in the literature but, unfortunately, there is no formal definition for such a concept, to the best of our knowledge. To cover this lack, we proceed to extend here the idea of unfounded sets from disjunctive logic programs to the epistemic case. For this purpose, we focus next on the original language of *epistemic specifications* [23] (a fragment of epistemic theories closer to logic programs) on which most approaches have been actually defined.

Let us start by introducing some terminology. An *objective literal* is either an atom $a \in At$, its negation $\neg a$ or its double negation $\neg\neg a$. A *subjective literal* is any of the formulas[5] $\mathbf{K}\, l$, $\neg \mathbf{K}\, l$ or $\neg\neg \mathbf{K}\, l$ where l is an objective literal. A *literal* is either an objective or a subjective literal, and is called *negative* if it contains negation and *positive* otherwise. A *rule* is a formula of the form

$$a_1 \vee \ldots \vee a_n \leftarrow B_1 \wedge \ldots \wedge B_m \tag{5}$$

with $n \geq 0$, $m \geq 0$ and $m+n > 0$, where each a_i is an atom and each B_j is a literal. For any rule r like (5), we define its body as $\mathit{Body}(r) \stackrel{\text{def}}{=} B_1 \wedge \ldots \wedge B_m$ and its head $\mathit{Head}(r) \stackrel{\text{def}}{=} a_1 \vee \ldots \vee a_n$, which we sometimes use as the set of atoms $\{a_1, \ldots, a_n\}$. We define the subset $\mathit{Body}_{ob}(r)$ as those atoms occurring in objective literals in the body. Similarly, $\mathit{Body}_{sub}(r)$ is the set of atoms occurring in subjective literals. Note that $\mathit{Body}(r) = \mathit{Body}_{ob}(r) \cup \mathit{Body}_{sub}(r)$

[5]We focus here on the study of operator $\mathbf{K}$, but epistemic specifications also allow a second operator $\mathbf{M}$ whose relation to $\mathbf{K}$ is also under debate, so we leave its study for future work.

but $Body_{ob}(r) \cap Body_{sub}(r)$ is not necessarily empty. Moreover, these sets are restricted into the respective $Body^{+}_{ob}(r)$ and $Body^{+}_{sub}(r)$ that further require those atoms to occur in positive body literals.

When $n = 0$, $Head(r) = \perp$ and the rule is a *constraint*, whereas if $m = 0$ then $Body(r) = \top$ and the rule is a *fact*.

An *epistemic specification* or *program* is a set of rules. As with formulas, a program without occurrences of $\mathbf{K}$ is said to be objective (it corresponds to a standard disjunctive logic program with double negation). For this case, and to allow a better comparison, we reproduce below the standard definition [36, Definition 3.1] of *unfounded set* (we call it "objective" to distinguish it from our extension).

Definition 3 (Objective unfounded set [36])**.** *Let* Π *be an objective program and* I *a propositional interpretation. A set of atoms* X *is an* (objective) unfounded set *with respect to* Π *and* I *if there is no rule* $r \in \Pi$ *with* $Head(r) \cap X \neq \emptyset$ *satisfying:*

1. $I \models Body(r)$
2. $Body^{+}_{ob}(r) \cap X = \emptyset$
3. $(Head(r) \setminus X) \cap I = \emptyset$ □

Now, we introduce the following extension for epistemic programs:

Definition 4 (Unfounded set)**.** *Let* Π *be a program and* $\mathbb{W}$ *a belief view. An* unfounded set $\mathbb{S}$ *with respect to* Π *and* $\mathbb{W}$ *is a non-empty set of pairs where, for each* $\langle X, I\rangle \in \mathbb{S}$*, it follows that* X *and* I *are sets of atoms and there is no rule* $r \in \Pi$ *with* $Head(r) \cap X \neq \emptyset$ *satisfying:*

1. $(\mathbb{W}, I) \models Body(r)$
2. $Body^{+}_{ob}(r) \cap X = \emptyset$
3. $(Head(r) \setminus X) \cap I = \emptyset$
4. *For all* $\langle X', I'\rangle \in \mathbb{S}$*,* $Body^{+}_{sub}(r) \cap X' = \emptyset$ □

As we can see, we added a fourth condition, but the first three ones essentially preserve Definition 3 except that we use $(\mathbb{W}, I)$ to check satisfaction of $Body(r)$, as it may contain now subjective literals. Intuitively, each I represents some potential belief set (or stable model) and X is some set of atoms without a "justifying" rule, that is, there is no $r \in \Pi$ allowing a positive derivation of atoms in X. A rule like that should have a true $Body(r)$ (condition 1) but not because of positive literals in X (condition 2) and is not used to derive other head atoms outside X (condition 3). The novelty in our definition is the addition of condition 4: to consider r a justifying rule, we additionally require not using any positive literal $\mathbf{K}\, a$ in the body such that atom a also belongs to any of the unfounded components X' in $\mathbb{S}$.

Definition 5 (Founded world view)**.** *Let Π be a program and $\mathbb{W}$ be a belief view. We say that $\mathbb{W}$ is* unfounded *if there is some unfounded-set $\mathbb{S}$ such that every $\langle X, I\rangle \in \mathbb{S}$ satisfies $I \in \mathbb{W}$ and $X \cap I \neq \emptyset$. $\mathbb{W}$ is called* founded *otherwise.* □

When Π is an objective program, each pair $\langle X, I\rangle$ corresponds to a standard unfounded set X of some potential stable model I in the traditional sense of [36].

Example 4. *Given the single disjunctive rule $a \vee b$ suppose we check the (expected) world view $\mathbb{W} = [\{a\}, \{b\}]$. For $I = \{a\}$ and $X = \{a\}$, rule $a \vee b$ satisfies the four conditions and justifies a. The same happens for $I = \{b\} = X$. So, $\mathbb{W}$ is founded. However, suppose we try with $\mathbb{W}' = [\{a, b\}]$ instead. For $I = \{a, b\}$ we can form $X = \{a\}$ and $X' = \{b\}$ and in both cases, the only rule in the program, $a \vee b$, violates condition 3. As a result, $\mathbb{W}'$ is unfounded due to the set $\mathbb{S}' = \{\langle\{b\}, \{a, b\}\rangle, \langle\{a\}, \{a, b\}\rangle\}$.* □

To illustrate how condition 4 works, let us continue with Example 3.

Example 5 (Example 3 continued)**.** *Theory $\Gamma_3 = \{\mathbf{K}\, a \to a\}$ is also a program. Given belief set $\mathbb{W} = [\{a\}]$ we can observe that $\mathbb{S} = [\langle\{a\}, \{a\}\rangle]$ makes $\mathbb{W}$ unfounded because the unique rule in Γ_3 does not fulfil condition 4: we cannot derive atom a from a rule that contains $a \in Body^+_{sub}(r)$. On the other hand, the other G94 world view, $\mathbb{W} = [\emptyset]$, is trivially founded.* □

Since Definition 5 only depends on some epistemic program and its selected world views, we can raise it to a general property for any epistemic semantics.

Property 1 (Foundedness)**.** *A semantics satisfies* foundedness *when all the world views it assigns to any program Π are founded.* □

An interesting observation is that in all the original examples of epistemic specifications [23, 26] used by Gelfond to introduce G94, modal operators occurred in the scope of negation. Since unfoundedness is never raised by negative subjective literals, unfounded world views could not be spotted using this family of examples.

5 Foundedness in the previously existing semantics

As we discussed in Example 5, our introduced definition of *unfounded world view* allowed disregarding the self-supported solution $[\{a\}]$ for program $\{\mathbf{K}\, a \to a\}$ obtained by G94. This immediately implies that G94 does not satisfy foundedness, as expected. In fact, this is not surprising, since all the remaining approaches previously existing in the literature were precisely proposed to disregard (among other cases) world view $[\{a\}]$ for that example. What is more striking, however, is that none of those approaches actually satisfy foundedness, as we proceed to prove next for each of them through counterexamples.

Foundedness in the G11 and K15 semantics

As mentioned in the introduction, Gelfond [24] and, later, Kahl et al. [34] revisited the semantics of epistemic logic programs in an attempt to get rid of unintended world views. These approaches, which we respectively denote as G11 and K15, were based on the same idea: modifying the definition of subjective reduct $\Gamma^{\mathbb{W}}$ so that some occurrences of subjective literals $\mathbf{K}\, l$ are not replaced by a truth constant, but by the objective literal l instead.

Definition 6 (G11-world views)**.** *Given a logic program* Π*, its G11-reduct with respect to a non-empty set of interpretations* $\mathbb{W}$ *is obtained by:*

1. *replacing by* $\bot$ *every subjective literal* $L \in Body_{sub}(r)$ *such that* $\mathbb{W} \not\models L$*,*
2. *removing all other occurrences of subjective literals of the form* $\neg\, \mathbf{K}\, l$*,*
3. *replacing all other occurrences of subjective literals of the form* $\mathbf{K}\, l$ *by* l*.*

A non-empty set of interpretations $\mathbb{W}$ *is a G11-world view of* Π *iff* $\mathbb{W}$ *is the set of all stable models of the G11-reduct of* Π *with respect to* $\mathbb{W}$*.* $\square$

Then, we can use the following example to show that this semantics does not satisfy foundedness:

Counterexample 1. *Take the epistemic logic program:*

$$a \vee b \qquad\qquad a \leftarrow \mathbf{K}\, b \qquad\qquad b \leftarrow \mathbf{K}\, a \qquad\qquad (\Pi_1)$$

whose G94-world views are $\mathbb{W} = [\{a\}, \{b\}]$ *and* $\mathbb{W}' = [\{a, b\}]$*. These are, indeed, the two cases we analysed in Example 4.* $\mathbb{W}$ *is again founded because* $a \vee b$ *keeps justifying both possible* $\langle X, I\rangle$ *pairs, that is,* $[\langle\{a\}, \{a\}\rangle]$ *and* $[\langle\{b\}, \{b\}\rangle]$*. However, for the world view* $\mathbb{W}'$*, we still have the unfounded set* $\mathbb{S}' = [\langle\{a\}, \{a, b\}\rangle, \langle\{b\}, \{a, b\}\rangle]$ *which violates condition 3 for the first rule as before, but also condition 4 for the other two rules.* □

Note how $\mathbb{S}'$ allows us to spot the root of the derivability problem: to justify a in $\langle\{a\}, \{a, b\}\rangle$ we cannot use $a \leftarrow \mathbf{K}\, b$ because b is part of the unfounded structure X in the other pair $\langle\{b\}, \{a, b\}\rangle$, and vice versa. It is not difficult to see that this unfounded G94-world view is also a G11-world view, so G11 *does not satisfy foundedness* either. To see why, note that the G11-reduct of program Π_1 with respect to $\mathbb{W}'$ is the objective program:

$$a \vee b \qquad\qquad a \leftarrow b \qquad\qquad b \leftarrow a$$

which has the unique stable model $\{a, b\}$ which is the only element in $\mathbb{W}'$.

The K15-reduct is a slight variant of the G11-reduct, so that $\mathbf{K}\, l$ is always replaced by l in any subjective literal satisfied by $\mathbb{W}$, even if $\mathbf{K}\, l$ is in the scope of negation.

Definition 7 (K15-world views)**.** *Given a logic program* Π*, its K15-reduct with respect to a non-empty set of interpretations* $\mathbb{W}$ *is obtained by:*

1. *replacing by* $\bot$ *every subjective literal* $L \in Body_{sub}(r)$ *such that* $\mathbb{W} \not\models L$*,*
2. *replacing all other occurrences of subjective literals of the form* $\mathbf{K}\, l$ *by* l*.*

A non-empty set of interpretations $\mathbb{W}$ *is a K15-world view of* Π *iff* $\mathbb{W}$ *is the set of all stable models of the K15-reduct of* Π *with respect to* $\mathbb{W}$*.* □

In general, the K15 and G11 reducts may differ. Yet, for Counterexample 1, it is easy to see that the K15-reduct of program Π_1 with respect to $\mathbb{W}'$ is the same as its G11-reduct and, thus, we get that $\mathbb{W}'$ is also an unfounded K15-world view of Π_1.

Foundedness in the F15 and S17 semantics

A more elaborated strategy is adopted by the recent approaches by Fariñas et al. [21, 58] (F15) and Shen and Eiter [56] (S17), that treat the previous world views as candidate solutions[6], but select the ones with *minimal knowledge* in a second step. This allows removing the unfounded world view $[\{a, b\}]$ in Counterexample 1, because the other solution $[\{a\}, \{b\}]$ provides less knowledge. Unfortunately, this strategy does not suffice to guarantee foundedness, since other formulas (such as constraints) may remove the founded world view without providing justification for the unfounded one, as explained below.

Let us recall the F15 semantics, based on a combination of Equilibrium Logic with the modal logic S5. We follow here the revision made in [58].

Definition 8. *An EHT-interpretation is a pair $\langle \mathbb{W}, h\rangle$ where $\mathbb{W}$ is a non-empty set of interpretations and $h : \mathbb{W} \longrightarrow 2^{At}$ is a function mapping each interpretation T to some subset of atoms such that $h(T) \subseteq T$.* □

Satisfaction of formulas with respect to EHT-interpretations is defined in a similar way as with respect to belief interpretations. Satisfaction of a formula φ with respect to an EHT-interpretation $\langle \mathbb{W}, h\rangle$ and a point $I \in \mathbb{W}$ is recursively defined as follows:

1. $\langle \mathbb{W}, h, I\rangle \not\models \bot$,
2. $\langle \mathbb{W}, h, I\rangle \models a$ iff $a \in I$, for any atom $a \in At$,
3. $\langle \mathbb{W}, h, I\rangle \models \psi_1 \wedge \psi_2$ iff $\langle \mathbb{W}, h, I\rangle \models \psi_1$ and $\langle \mathbb{W}, h, I\rangle \models \psi_2$,
4. $\langle \mathbb{W}, h, I\rangle \models \psi_1 \vee \psi_2$ iff $\langle \mathbb{W}, h, I\rangle \models \psi_1$ or $\langle \mathbb{W}, h, I\rangle \models \psi_2$,
5. $\langle \mathbb{W}, h, I\rangle \models \psi_1 \rightarrow \psi_2$ iff $\langle \mathbb{W}, h', I\rangle \not\models \psi_1$ or $\langle \mathbb{W}, h', I\rangle \models \psi_2$ for both $h' = h$ and $h' = id$, and
6. $\langle \mathbb{W}, h, I\rangle \models \mathbf{K}\,\psi$ iff $\langle \mathbb{W}, h, J\rangle \models \psi$ for all $J \in \mathbb{W}$.

where $id : \mathbb{W} \longrightarrow 2^{At}$ is the identity function mapping, that is, $id(T) = T$ for every $T \in \mathbb{W}$. We say that an EHT-interpretation $\langle \mathbb{W}, h\rangle$ *satisfies* a formula φ

[6] In [21, 58], these candidate world views are called *epistemic equilibrium models* while selected world views receive the name of *autoepistemic equilibrium models*.

when $\langle \mathbb{W}, h, I\rangle \models \varphi$ for all $I \in \mathbb{W}$. We say that $\langle \mathbb{W}, h\rangle$ *satisfies* a theory Γ, written $\langle \mathbb{W}, h\rangle \models \Gamma$, if it satisfies all its formulas $\varphi \in \Gamma$. In this last case, we also say that $\langle \mathbb{W}, h\rangle$ is an *EHT-model* of Γ.

An EHT-interpretation $\langle \mathbb{W}, h\rangle$ is said to be *total on* a set $\mathbb{X} \subseteq \mathbb{W}$ iff $h(I) = I$ for every $I \in \mathbb{X}$. It is said to be just *total* iff it is total on $\mathbb{W}$. Then, equilibrium models are defined as follows:

Definition 9. *A total EHT-model $\langle \mathbb{W}, id\rangle \models \Gamma$ of some theory Γ is an* equilibrium EHT-model *iff there is no other EHT-model $\langle \mathbb{W}, h\rangle \models \Gamma$ such that $h(I) \subset I$ for some $I \in \mathbb{W}$.* □

The F15-world views are obtained from a selection among equilibrium EHT-models. For defining that selection, we need to introduce the following definitions:

Definition 10. *Given a theory Γ, a non-empty sets of interpretations and a subset $\mathbb{X} \subseteq \mathbb{W}$, we write $\mathbb{W}, \mathbb{X} \models^* \Gamma$ iff the following two conditions are satisfied:*

1. *$\langle \mathbb{W}, id, I\rangle \models \Gamma$ for all $I \in \mathbb{X}$, and*
2. *if $\langle \mathbb{W}, h, I\rangle \models \Gamma$ for some $I \in \mathbb{W}$ such that $\langle \mathbb{W}, h\rangle$ is total on $\mathbb{W} \setminus \mathbb{X}$, then $\langle \mathbb{W}, h\rangle$ is total.* □

Then, for any two pairs of non-empty set of interpretations $\mathbb{W}$ and $\mathbb{W}'$ we write $\mathbb{W} \leq_\Gamma \mathbb{W}'$ iff

$$\mathbb{W} \cup \{I\}, \mathbb{W} \models^* \Gamma \quad \textit{implies} \quad \mathbb{W}' \cup \{I\}, \mathbb{W}' \models^* \Gamma$$

for every I such that I belongs to some equilibrium EHT-model of Γ. As usual $\mathbb{W} <_\Gamma \mathbb{W}'$ stands for $\mathbb{W} \leq_\Gamma \mathbb{W}'$ and $\mathbb{W} \not\leq_\Gamma \mathbb{W}'$. □

Definition 11. *Given a theory Γ, an equilibrium EHT-model $\mathbb{W}$ is called an* F15-world view *iff there is no other equilibrium EHT-model $\mathbb{W}'$ such that $\mathbb{W} \subset \mathbb{W}'$ or $\mathbb{W} <_\Gamma \mathbb{W}'$.* □

Note that F15-world views are chosen as a kind of minimal equilibrium EHT-models. In many examples in the literature, those minimal models happen to be founded, leading to the wrong impression that foundedness is satisfied. However, in those same examples, some non-minimal equilibrium EHT-models are, in fact, unfounded. Given that F15 allows ruling out a specific equilibrium EHT-model by adding a constraint, we can easily force the unfounded world views to come out as selected models. This is shown in the following counterexample.

Counterexample 2 (Counterexample 1 continued)**.** *Take now the program* $\Pi_2 = \Pi_1 \cup \{\bot \leftarrow \neg\mathbf{K}\, a\}$*. This new subjective constraint rules out world view* $\mathbb{W} = [\{a\}, \{b\}]$ *because the latter satisfies* $\neg\mathbf{K}\, a$*. Then, the G94, G11 and K15 semantics assign* $\mathbb{W}' = [\{a, b\}]$ *as the unique world view of this program. Similarly,* $\mathbb{W}'$ *is also its unique equilibrium EHT-model and, thus, an F15-world view. However,* $\mathbb{W}'$ *is still unfounded in* Π_2 *because constraints do not affect that feature (their empty head never justifies any atom).* □

Counterexample 2 not only shows that F15 does not satisfy foundedness, but also that the strategy of ruling out unfounded world views by relying on a minimisation is doomed to failure if we can remove candidate world views using constraints or other kind of rules. This is, in fact, the problem we can also find in the S17 semantics [56]. Recall that S17 uses a different syntax to the rest of approaches: instead of the epistemic operator $\mathbf{K}$, it uses the *epistemic negation* operator **not**. However, Son et al. [57] showed that '**not** l' can be defined in terms of $\mathbf{K}$ as '*not* $\mathbf{K}\, l$' and, then, characterised the S17-world views as a class of minimal K15-world views.

Definition 12 (S17-world views)**.** *Let* Π *be a logic program* Π *and* E_Π *be the set of epistemic literals that contains not* $\mathbf{K}\, l$ *for every epistemic literal of the form* $\mathbf{K}\, l$ *that occurs in* Π*. Let* $\Phi_{\mathbb{W}} \stackrel{\text{def}}{=} \{\ L \in E_\Pi \mid \mathbb{W} \models L\ \}$ *be the subset of* E_Π *satisfied by the non-empty set of interpretations* $\mathbb{W}$*. Then, a non-empty set of interpretations* $\mathbb{W}$ *is an S17-world view iff it is a K15-world view and there is no other K15-world view* $\mathbb{W}'$ *such that* $\Phi_{\mathbb{W}'} \supset \Phi_{\mathbb{W}}$*.* □

Using this characterisation, we can use again Counterexample 2 to prove that S17 does not satisfy foundedness, since the program Π_2 had a *unique* K15-world view $\mathbb{W} = [\{a, b\}]$ that happened to be unfounded. Being $\mathbb{W}$ the unique K15 world view, the additional minimisation imposed by S17 makes no difference with respect to K15 in this case.

6 Founded Autoepistemic Equilibrium Logic

We present now the semantics proposed in this paper, introducing *Founded Autoepistemic Equilibrium Logic* (FAEEL). As suggested by the similarity in their names, FAEEL follows the same spirit as F15, that is, it combines Equilibrium Logic with a modal approach, but replaces S5 by Moore's Autoepistemic Logic (AEL). Note that this implies combining two non-monotonic formalisms, since AEL is non-monotonic too[7]. We will begin defining a combination of the monotonic bases of equilibrium logic and AEL: the intermediate logic of HT and the modal logic KD45, respectively.

We start by introducing an elaboration of the belief (or KD45) interpretation $(\mathbb{W}, I)$ already seen but replacing belief sets by HT pairs. Thus, we extend now the idea of *belief view* $\mathbb{W}$ to a non-empty set of HT-interpretations $\mathbb{W} = \{\langle H_1, T_1\rangle, \ldots, \langle H_n, T_n\rangle\}$ and say that $\mathbb{W}$ is *total* when $H_i = T_i$ for all of them, coinciding with the form of belief views $\mathbb{W} = \{T_1, \ldots, T_n\}$ we had so far. Given an arbitrary $\mathbb{W} = \{\langle H_1, T_1\rangle, \ldots, \langle H_n, T_n\rangle\}$ we define its corresponding total belief view as $\mathbb{W}^t \stackrel{\text{def}}{=} \{T_1, \ldots, T_n\}$. A *belief interpretation* $\mathcal{I}$ is now redefined as $\mathcal{I} = (\mathbb{W}, \langle H, T\rangle)$, or simply $\mathcal{I} = (\mathbb{W}, H, T)$, where $\mathbb{W}$ is a belief view and $\langle H, T\rangle$ stands for the real world, possibly not in $\mathbb{W}$. Therefore, $\mathcal{I}$ consists of a belief view $\mathbb{W}$ and a real world $\langle H, T\rangle$ where we may have partial or total information in any of them now. When $\mathbb{W}$ is total, that is $\mathbb{W}^t = \mathbb{W}$, we say that $\mathcal{I}$ is a *total-view* interpretation and further say that $\mathcal{I}$ is (completely) *total* if, additionally, $H = T$. Next, we redefine the satisfaction relation as follows. A belief interpretation $\mathcal{I} = (\mathbb{W}, H, T)$ satisfies a formula φ, written $\mathcal{I} \models \varphi$, iff:

- $\mathcal{I} \not\models \bot$,
- $\mathcal{I} \models a$ iff $a \in H$, for any atom $a \in At$,
- $\mathcal{I} \models \psi_1 \wedge \psi_2$ iff $\mathcal{I} \models \psi_1$ and $\mathcal{I} \models \psi_2$,
- $\mathcal{I} \models \psi_1 \vee \psi_2$ iff $\mathcal{I} \models \psi_1$ or $\mathcal{I} \models \psi_2$,
- $\mathcal{I} \models \psi_1 \rightarrow \psi_2$ iff both:
  (i) $\mathcal{I} \not\models \psi_1$ or $\mathcal{I} \models \psi_2$; and
  (ii) $(\mathbb{W}^t, T) \not\models \psi_1$ or $(\mathbb{W}^t, T) \models \psi_2$.
- $\mathcal{I} \models \mathbf{K}\,\psi$ iff $(\mathbb{W}, H_i, T_i) \models \psi$ for all $\langle H_i, T_i\rangle \in \mathbb{W}$.

For total belief interpretations, this new satisfaction relation collapses to the modal logic KD45. For this reason, we compactly write total belief interpretations just as $(\mathbb{W}, T)$ instead of $(\mathbb{W}, T, T)$. Interpretation $(\mathbb{W}, H, T)$ is a *belief model* of Γ iff $(\mathbb{W}, H_i, T_i) \models \varphi$ for all $\langle H_i, T_i\rangle \in \mathbb{W} \cup \{\langle H, T\rangle\}$ and all $\varphi \in \Gamma$ – additionally, when $\langle H, T\rangle \in \mathbb{W}$, we further say that $\mathbb{W}$ is an *epistemic model*

[7]It is perhaps worth to remember that non-monotonic S5 collapses to S5, so this double non-monotonicity did not arise in F15.

of Γ, abbreviated as $\mathbb{W} \models \Gamma$. Given any epistemic theory Γ, we can force its models to be total-view by adding the axiom schemata:

$$\mathbf{K}(a \vee \neg a) \tag{6}$$

for every atom $a \in At$. On the other hand, the addition of (4) forces total belief models (i.e., both the belief view and the real world are total). This is formally stated below:

Proposition 2. *Let Γ be an epistemic theory for signature At. Then:*

- *$\mathcal{I}$ is a model of $\Gamma \cup (6)$ iff both $\mathcal{I}$ is a model of Γ and $\mathcal{I}$ is total-view.*
- *$\mathcal{I}$ is a model of $\Gamma \cup (4)$ iff both $\mathcal{I}$ is a model of Γ and $\mathcal{I}$ is total.* □

Note how axiom (4) is, in fact, stronger than (6) since any total belief model is also total-view.

A fundamental property from intuitionistic logic related to our total interpretations is *persistence*. Informally speaking, persistence means that any formula satisfied in a point of a Kripke structure is also preserved in all its accessible points. In the case of the intermediate logic of HT, this means that anything true "here" should also hold "there." In FAEEL this means that any formula satisfied by a belief interpretation must also be satisfied by its corresponding total interpretation, as stated next.

Proposition 3 (Persistence)**.** $(\mathbb{W}, H, T) \models \varphi$ *implies* $(\mathbb{W}^t, T) \models \varphi$. □

Proof. Just note that, for any atom a, it follows that $\langle \mathbb{W}, H, T \rangle \models a$ iff $a \in H \subseteq T$ which implies $\langle \mathbb{W}^t, T \rangle \models a$. The rest of the proof follows by induction on the structure of φ. □

A belief model just captures collections of HT models which need not be in equilibrium. To make the agent's beliefs correspond to stable models we impose a particular minimisation criterion. We begin defining an ordering relation among belief views as follows.

Definition 13. *Given belief views $\mathbb{W}'$ and $\mathbb{W}$, we write $\mathbb{W}' \preceq \mathbb{W}$ iff the following two conditions hold:*

(i) for every $\langle H, T \rangle \in \mathbb{W}$, there is some $\langle H', T \rangle \in \mathbb{W}'$, with $H' \subseteq H$.

(ii) for every $\langle H', T \rangle \in \mathbb{W}'$, there is some $\langle H, T \rangle \in \mathbb{W}$, with $H' \subseteq H$.

As usual, we write $\mathbb{W}' \prec \mathbb{W}$ *iff* $\mathbb{W}' \preceq \mathbb{W}$ *and* $\mathbb{W}' \neq \mathbb{W}$. □

This ordering relation is extended to any pair of belief interpretations $\mathcal{I}' = (\mathbb{W}', H', T')$ and $\mathcal{I} = (\mathbb{W}, H, T)$ so that we write $\mathcal{I}' \preceq \mathcal{I}$ when $T' = T$, $H' \subseteq H$ and $\mathbb{W}' \preceq \mathbb{W}$. Again, $\mathcal{I}' \prec \mathcal{I}$ means $\mathcal{I}' \preceq \mathcal{I}$ and $\mathcal{I}' \neq \mathcal{I}$. The intuition for $\mathcal{I}' \preceq \mathcal{I}$ is that $\mathcal{I}'$ contains less information than $\mathcal{I}$ for each fixed T_i component. As a result, $\mathcal{I}' \models \varphi$ implies $\mathcal{I} \models \varphi$ for any formula φ without implications other than $\neg\psi = \psi \to \bot$.

Definition 14. *A total belief interpretation* $\mathcal{I} = (\mathbb{W}, T)$ *is said to be an* equilibrium belief model *of some theory* Γ *iff* $\mathcal{I}$ *is a belief model of* Γ *and there is no other belief model* $\mathcal{I}'$ *of* Γ *such that* $\mathcal{I}' \prec \mathcal{I}$. □

By EQB[Γ] we denote the set of equilibrium belief models of Γ. As a final step, we impose a fixpoint condition to minimise the agent's knowledge as follows.

Definition 15. *A belief view* $\mathbb{W}$ *is called a* FAEEL-world view *of* Γ *iff:*

$$\mathbb{W} \quad = \quad \{\ T \mid (\mathbb{W}, T) \in \mathrm{EQB}[\Gamma]\ \}$$ □

The logic induced by equilibrium world views is called *Founded Autoepistemic Equilibrium Logic* (FAEEL).

Example 6 (Example 5 continued)**.** *Back to* $\Gamma_3 = \{\mathbf{K}\, a \to a\}$*, remember its unique founded G94-world view was* $[\emptyset]$*. It is easy to see that* $\mathcal{I} = ([\emptyset], \emptyset) \in \mathrm{EQB}[\Gamma_3]$ *because* $([\emptyset], \emptyset) \models \Gamma_3$ *and no smaller belief model can be obtained. Moreover,* $[\emptyset]$ *is an equilibrium world view of* Γ_3 *since no other* $T \notin [\emptyset]$ *satisfies* $([\emptyset], T) \in \mathrm{EQB}[\Gamma_3]$*. The only possibility is* $([\emptyset], \{a\})$ *but it fails because there is a smaller belief model* $([\emptyset], \emptyset, \{a\})$ *satisfying* $\mathbf{K}\, a \to a$*. As for the other potential world view* $[\{a\}]$*, it is not in equilibrium: we already have* $\mathcal{I}' = ([\{a\}], \{a\}) \notin \mathrm{EQB}[\Gamma_3]$ *because the smaller interpretation* $\mathcal{I}'' = ([\langle\emptyset, \{a\}\rangle], \{a\}, \{a\})$ *also satisfies* Γ_3*. In particular, note that* $\mathcal{I}'' \not\models \mathbf{K}\, a$ *and, thus, clearly satisfies* $\mathbf{K}\, a \to a$. □

In the rest of this section we explore the relation between FAEEL and G94. We know that some G94 world views are not FAEEL world views, as happens with $[\{a\}]$ above. Still, we will prove that the opposite *does hold*, that is, any FAEEL world view is always a G94 world view, so the former constitutes a *strictly stronger* semantics. The key point for that result is that

G94 world views of a theory Γ can be characterised as the equilibrium belief models of $\Gamma \cup (6)$ – remember that, by Proposition 2, this means using total-view models of Γ in the minimisation process. We begin observing that the addition of (6) to a theory Γ produces a superset of equilibrium models:

Proposition 4. *Given any epistemic theory* Γ*,* $\mathrm{EQB}[\Gamma] \subseteq \mathrm{EQB}[\Gamma \cup (6)]$. $\square$

Proof. By definition, $\mathcal{I} \in \mathrm{EQB}[\Gamma]$ iff $\mathcal{I}$ is a total model of Γ and there is no belief model $\mathcal{I}'$ of Γ such tat $\mathcal{I}' \prec \mathcal{I}$. Therefore, there is no total-view model $\mathcal{I}'$ of Γ with $\mathcal{I}' \prec \mathcal{I}$ either. But this is the same than saying that there is no model $\mathcal{I}'$ of $\Gamma \cup (6)$ with $\mathcal{I}' \prec \mathcal{I}$. Finally, as $\mathcal{I}$ is total, it also satisfies (6) and so, we conclude $\mathcal{I} \in \mathrm{EQB}[\Gamma \cup (6)]$. $\square$

Note that the opposite direction is not necessarily true. For instance, $([\{a\}], \{a\})$ is an equilibrium model of $\{\mathbf{K}\, a \to a\} \cup (6)$ but not of $\{\mathbf{K}\, a \to a\}$ alone. The following lemma relates FAEEL total-view interpretations forced by (6) with the epistemic reduct transformation $\varphi^{\mathbb{W}}$ used in G94 semantics.

Lemma 1. *Let* Γ *be a formula and* $\mathcal{I} = (\mathbb{W}, H, T)$ *be a total-view interpretation. Then,* $\mathcal{I}$ *is a model of* φ *iff* $\mathcal{I}$ *is a model of* $\varphi^{\mathbb{W}}$. $\square$

Proof. Assume that $\varphi = \mathbf{K}\, \psi$. Then, we get the equivalences:

$$\begin{array}{ll} & (\mathbb{W}, H', T') \models \varphi \\ \text{iff} & (\mathbb{W}, T'', T'') \models \psi \text{ for every } \langle T'', T'' \rangle \in \mathbb{W} \\ \text{iff} & \mathbb{W} \text{ is an S5-model of } \varphi \\ \text{iff} & \varphi^{\mathbb{W}} = \top \\ \text{iff} & \mathcal{I} \models \varphi^{\mathbb{W}} \end{array}$$

Then, by induction on the structure of φ, we conclude that $\mathcal{I} \models \varphi$ iff $\mathcal{I} \models \varphi^{\mathbb{W}}$. Finally, the following equivalences can be obtained:

$$\begin{array}{ll} & \mathcal{I} \text{ is a model of } \varphi \\ \text{iff} & \mathcal{I} \models \varphi \text{ and } (\mathbb{W}, T', T') \models \varphi \text{ for all } \langle T', T' \rangle \in \mathbb{W} \\ \text{iff} & \mathcal{I} \models \varphi^{\mathbb{W}} \text{ and } (\mathbb{W}, T', T') \models \varphi^{\mathbb{W}} \text{ for all } \langle T', T' \rangle \in \mathbb{W} \\ \text{iff} & \mathcal{I} \text{ is a model of } \varphi^{\mathbb{W}} \end{array}$$

$\square$

The result of the epistemic reduct $\varphi^{\mathbb{W}}$ is a propositional (i.e. objective) theory. The next two lemmata provide a pair of properties of propositional theories in our semantics.

Lemma 2. *Let Γ be a propositional theory and $\mathcal{I} = (\mathbb{W}, H, T)$ be some interpretation. Then, $\mathcal{I}$ is a model of Γ iff $\langle H', T' \rangle$ is an HT-model of Γ for every HT-interpretation $\langle H', T' \rangle \in \mathbb{W} \cup \{\langle H, T \rangle\}$.* $\square$

Proof. By definition, it follows that $\mathcal{I}$ is a model of Γ iff $\mathcal{I}$ is a model of φ for all $\varphi \in \Gamma$. Furthermore, $\mathcal{I}$ is a model of φ iff $\mathcal{I} \models \varphi$ and $(\mathbb{W}, H', T') \models \varphi$ for every $\langle H', T' \rangle \in \mathbb{W}$ iff $(\mathbb{W}, H', T') \models \varphi$ for every HT-interpretation $\langle H', T' \rangle \in \mathbb{W} \cup \{\langle H, T \rangle\}$. Finally, since φ is a propositional formula, it follows that $(\mathbb{W}, H', T') \models \varphi$ iff $\langle H', T' \rangle \models \varphi$. Hence, $\mathcal{I}$ is a model of Γ iff $\langle H', T' \rangle$ is an HT-model of Γ for every HT-interpretation $\langle H', T' \rangle \in \mathbb{W} \cup \{\langle H, T \rangle\}$. $\square$

Lemma 3. *Let Γ be a propositional theory and $\mathcal{I} = (\mathbb{W}, T)$ be some total interpretation. Then, $\mathcal{I}$ is a equilibrium model of Γ iff T' is an equilibrium model of Γ for every $T' \in \mathbb{W} \cup \{T\}$.* $\square$

Proof. First note that, since Γ is propositional, $\mathcal{I}$ is a model of Γ iff T' is a model of Γ for every $T' \in \mathbb{W} \cup \{T\}$. Hence, we can conclude that $\mathcal{I}$ is an equilibrium model of Γ iff there is no model $\mathcal{I}' = (\mathbb{W}', H, T)$ of Γ such that $\mathcal{I}' \prec \mathcal{I}$.

Suppose, for the sake of contradiction, that there is some $T' \in \mathbb{W} \cup \{T\}$ which is not an equilibrium model of Γ. Then, there is $\langle H', T' \rangle \models \Gamma$ such that $H' \subset T'$. Let $\mathcal{I}' = (\mathbb{W}, H', T')$ if $T' = T$ and $\mathcal{I}' = (\mathbb{W}', T', T')$ with $\mathbb{W}' = \{\langle H', T' \rangle\} \cup \mathbb{W}$ otherwise. Then, $\mathcal{I}' \prec \mathcal{I}$ and, from Lemma 2, it follows that $\mathcal{I}'$ is a model of Γ which is a contradiction. Hence, T' must be an equilibrium model of Γ for every $T' \in \mathbb{W} \cup \{T\}$. The other way around, assume that T' is an equilibrium model of Γ for every $T' \in \mathbb{W} \cup \{T\}$ and suppose, for the sake of contradiction, that $\mathcal{I}$ is not a equilibrium model of Γ. Then, there is some model $\mathcal{I}' = (\mathbb{W}', H, T)$ of Γ such that $\mathcal{I}' \prec \mathcal{I}$. Then, from Lemma 2, we get that $\langle H', T' \rangle \models \Gamma$ for every $\langle H', T' \rangle \in \mathbb{W} \cup \{\langle H, T \rangle\}$. This implies that T' is not an equilibrium model of Γ which is a contradiction. Consequently, $\mathcal{I}$ must be a equilibrium model of Γ. $\square$

We are now ready to prove the characterisation of G94 world views of Γ in terms of equilibrium belief models of $\Gamma \cup (6)$.

Theorem 1. *The G94-world views of any theory Γ coincide precisely with the FAEEL-world views of $\Gamma \cup (6)$.* $\square$

Proof. From Lemmas 1 and 3, we can see that $(\mathbb{W}, T) \in \mathrm{EQB}[\Gamma \cup (6)]$ iff $(\mathbb{W}, T) \in \mathrm{EQB}[\Gamma^{\mathbb{W}} \cup (6)]$ iff $\mathbb{W} \cup \{T\} \subseteq \mathrm{SM}[\Gamma^{\mathbb{W}}]$. Furthermore, from Definition 15, we get that $\mathbb{W}$ is a FAEEL-world view of $\Gamma \cup (6)$ iff

$$\mathbb{W} \quad = \quad \{\ T \mid \langle \mathbb{W}, T \rangle \in \mathrm{EQB}[\Gamma \cup (6)]\ \}$$

which, as we just saw, can be rewritten as

$$\mathbb{W} \quad = \quad \{\ T \mid \mathbb{W} \cup \{T\} \subseteq \mathrm{SM}[\Gamma^{\mathbb{W}}]\ \}$$

which holds iff $\mathbb{W} = \mathrm{SM}[\Gamma^{\mathbb{W}}]$. That is, iff $\mathbb{W}$ is a G94-world view of Γ. $\square$

Now, to prove that FAEEL is stronger than G94, it suffices to use the above characterisation of the latter. In other words, we will show that any FAEEL-world view of Γ is also an equilibrium model of $\Gamma \cup (6)$.

Proposition 5. *If $\mathbb{W}$ is a FAEEL-world view of Γ then $\mathbb{W}$ is a FAEEL-world view of $\Gamma \cup (6)$.* $\square$

Proof. By Proposition 4, we know $\mathrm{EQB}[\Gamma] \subseteq \mathrm{EQB}[\Gamma \cup (6)]$ so we only need to prove that if $\mathbb{W}$ is a FAEEL-world view then there is no propositional interpretation of $T \notin \mathbb{W}$ such that $(\mathbb{W}, T) \in \mathrm{EQB}[\Gamma \cup (6)]$.

Suppose, for the sake of contradiction, that the opposite holds, i.e., there is some $T \notin \mathbb{W}$ such that $(\mathbb{W}, T) \in \mathrm{EQB}[\Gamma \cup (6)]$. Since $\mathbb{W}$ is a FAEEL-world view, this implies that $(\mathbb{W}, T)$ is not an equilibrium model of Γ and, thus, that there is some non-total-view model $\mathcal{I}' = (\mathbb{W}', H, T)$ of Γ such that $\mathcal{I}' \prec \mathcal{I}$. Hence, there is some $\langle H', T' \rangle \in \mathbb{W}'$ such that $H' \subset T'$. Let $\mathcal{I}'' = \langle \mathbb{W}', H', T' \rangle$. Then, it follows that $\mathcal{I}'' \prec (\mathbb{W}, T')$ and, since $\mathbb{W} \in \mathrm{EQB}[\Gamma]$ and $T \in \mathbb{W}$, it follows that $(\mathbb{W}, T')$ is an equilibrium model of Γ. These two facts together imply that $\mathcal{I}''$ is not a model of Γ. Therefore, there is a formula $\varphi \in \Gamma$ such that $\mathcal{I}''$ is not a model of φ and, thus, there is $\langle H'', T'' \rangle \in \mathbb{W}' \cup \{\langle H', T' \rangle\}$ such that $(\mathbb{W}', H'', T'') \not\models \varphi$. On the other hand, since $\mathcal{I}'$ is a model of Γ, we can see that $(\mathbb{W}', H''', T''') \not\models \varphi$ for every $\langle H''', T''' \rangle \in \mathbb{W}' \cup \{\langle H, T \rangle\}$. Thus, we conclude $H'' = H'$, $T'' = T'$ and $(\mathbb{W}', H', T') \not\models \varphi$. However, since $\langle H', T' \rangle \in \mathbb{W}'$, this implies that $\mathcal{I}' = (\mathbb{W}', H, T) \not\models \varphi$, which is a contradiction with the fact that $\mathcal{I}'$ is a model of Γ. Consequently, $\mathbb{W}$ is a FAEEL-world view of $\Gamma \cup (6)$. $\square$

Finally, this immediately allows us to conclude:

Theorem 2. *If $\mathbb{W}$ is a FAEEL-world view of Γ then $\mathbb{W}$ is also a G94-world view of Γ.*

Proof. Suppose $\mathbb{W}$ is a FAEEL-world view of Γ. From Proposition 5, we get that $\mathbb{W}$ is a FAEEL-world view of $\Gamma \cup (6)$ and, from Theorem 1, the latter means that $\mathbb{W}$ is a G94-world view of Γ. $\square$

Therefore, as we had foreseen, FAEEL is strictly stronger than G94. This is, in fact, a distinctive feature of our semantics that does not hold in other approaches in the literature, as proved by the following example.

Example 7. *The following program:*

$$a \vee b \qquad\qquad c \leftarrow \mathbf{K}\, a \qquad\qquad \bot \leftarrow \neg c \tag{Π_3}$$

has no G94-world views, but according to G11, K15, F15 and S17 has world view $[\{a, c\}]$*. This example was also used in [6] to show that these semantics do not satisfy another property, called there* epistemic splitting. $\square$

One approach in the literature related to G94 that deserves a special mention is the definition of *epistemic views* of a theory Γ introduced in [62], since their definition shares some formal similarities with $\mathrm{EQB}[\Gamma \cup (6)]$ and seems to provide the same results for many examples. Still, the following example shows that this does not hold in general, that is, epistemic views from [62] are different from $\mathrm{EQB}[\Gamma \cup (6)]$, i.e., different from G94-world views.

Example 8. *Consider the singleton theory* $\Gamma = \{\neg\neg a \wedge \mathbf{K}\, \varphi \to a\}$ *with* φ *the following formula* $\varphi = \neg\neg a \to a$*. Then,* $\mathbb{W} = [\emptyset, \{a\}]$ *is both a G94 and FAEEL-world view of* Γ*, but not an epistemic view. To see that* $\mathbb{W}$ *is a G94-world view of* Γ*, note that* $\Gamma^{\mathbb{W}} = \{\neg\neg a \wedge \top \to a\} \equiv \{\neg\neg a \to a\}$ *which has two stable models:* $\emptyset$ *and* $\{a\}$*. Furthermore, from Theorem 1 this implies that* $\mathbb{W} \in \mathrm{EQB}[\Gamma \cup (6)]$*. However,* $\mathbb{W}$ *is not an epistemic view because* $(\mathbb{W}, \emptyset, \{a\})$ *is a model of* Γ *in the sense of [62]. Note that* $(\mathbb{W}, \emptyset, \{a\}) \not\models \varphi$ *and that both* $\emptyset$ *and* $\{a\}$ *belong to* $\mathbb{W}$*. This implies that* $(\mathbb{W}, \emptyset, \{a\}) \not\models \mathbf{K}\, \varphi$ *and, thus, that* $(\mathbb{W}, \emptyset, \{a\})$ *is a model of* Γ*. On the other hand, in our logic* $(\mathbb{W}, \emptyset, \{a\}) \not\models \varphi$ *does not imply* $(\mathbb{W}, \emptyset, \{a\}) \not\models \mathbf{K}\, \varphi$*. In fact* $(\mathbb{W}, \emptyset, \{a\}) \models \mathbf{K}\, \varphi$ *holds because both* $(\mathbb{W}, \emptyset, \emptyset) \models \varphi$ *and* $(\mathbb{W}, \{a\}, \{a\}) \models \varphi$ *hold.* $\square$

To conclude this section, we recall a recent result from [19] that provides a quite general, sufficient syntactic condition on epistemic logic programs under which FAEEL and G94 coincide. In particular, this happens when there is a kind of acyclicity among positive subjective literals, revealing that the difference between FAEEL and G94 is related to the treatment of cycles, as we will confirm in the next section.

Formally, the positive epistemic dependence relation among atoms in a program Π is defined so that $dep^+(a, b)$ is true iff there is any rule $r \in \Pi$ such that $a \in \mathit{Head}(r) \cup \mathit{Body}_{ob}(r)$ and $b \in \mathit{Body}^+_{sub}(r)$.

Definition 16 (Epistemically tight program)**.** *We say that an epistemic program Π is* epistemically tight *if we can assign an integer mapping $\lambda : At \to \mathbb{N}$ to each atom such that*

(i) $\lambda(a) = \lambda(b)$ *for any rule $r \in \Pi$ and atoms $a, b \in (\mathit{Atoms}(r) \backslash \mathit{Body}_{sub}(r))$,*

(ii) $\lambda(a) > \lambda(b)$ *for any pair of atoms a, b satisfying $dep^+(a, b)$.* □

Theorem 3 (Theorem 8 in [19])**.** *FAEEL and G94-world views coincide for epistemically tight programs.* □

Notice that, when a program has all its subjective literals in the scope of negation, it is trivially tight (we can assign the same level to all atoms) and so FAEEL and G94 coincide. This is interesting since, as we discussed before, any semantics is also trivially founded for this kind of programs.

7 Characterisation as Founded G94-World Views

In this section, we review an alternative characterisation of FAEEL introduced by Fandinno [19] that will allow us to prove the main result in this paper, namely, that FAEEL precisely obtains those G94-world views that are founded. According to this alternative characterisation, FAEEL-word views are those G94-word views that are equilibrium models on a new logic we will call S5-Equilibrium Logic (or S5-EL). This logic is similar to FAEEL, but without the "autoepistemic" minimisation of knowledge. Technically, this makes things simpler in two distinct ways: (i) it directly uses belief views instead of belief interpretations and, as a result, (ii) it lacks the autoepistemic fixpoint condition (Definition 15).

Definition 17. *A total epistemic model $\mathbb{W}$ of a theory Γ is said to be an* S5-world view *iff there is no other epistemic model $\mathbb{W}'$ of Γ s.t. $\mathbb{W}' \prec \mathbb{W}$.* □

The following result from [19] shows that FAEEL-world views can be characterised in terms of G94 and S5-world views.

Theorem 4 (Theorem 4 from [19])**.** *For any theory Γ, a belief view $\mathbb{W}$ is a FAEEL-world view iff both (i) $\mathbb{W}$ is a G94-world view and (ii) $\mathbb{W}$ is an S5-world view.* □

Example 9 (Example 1 continued)**.** *Take again program* $\Pi_1 = \{(a \vee b), (a \leftarrow \mathbf{K}\, b), (b \leftarrow \mathbf{K}\, a)\}$ *whose G94-world views were* $\mathbb{W} = [\{a\}, \{b\}]$ *and* $\mathbb{W}' = [\{a, b\}]$*, the latter being unfounded. We can see first that* $\mathbb{W}$ *is an S5-world view, since we cannot build any smaller epistemic model* $\mathbb{W}'' \prec \mathbb{W}$ *because removing any atom from* $\{a\}$ *or* $\{b\}$ *would make formula* $a \vee b$ *in* Π_1 *unsatisfied. To see that* $\mathbb{W}'$ *is not an S5-world view, note that* $\mathbb{W}''' = \{\langle\{a\}, \{a, b\}\rangle, \langle\{b\}, \{a, b\}\rangle\}$ *is strictly smaller* $\mathbb{W}''' \prec \mathbb{W}$ *but is also an epistemic model of* Π_1*. This is because* $a \vee b$ *holds in all these worlds whereas, for the two implications, at the "here" level, neither* $\mathbf{K}\, a$ *or* $\mathbf{K}\, b$ *hold, so implications are trivially true, whereas at the "there" level* $\mathbf{K}\, a$*,* $\mathbf{K}\, b$*,* a *and* b *hold. From Theorem 4 we conclude that* $\mathbb{W}$ *is the only FAEEL-world view of* Π_1*.* □

The interest of this alternative characterisation is that, for the syntax of epistemic programs, there is a strong connection between S5-EL and the foundedness condition. In particular, we will begin proving that S5-world views are always founded which, due to Theorem 4 above, will immediately mean that FAEEL satisfies foundedness too. Our proof follows a similar structure as the one for standard Equilibrium Logic [16]. Let us start by defining some auxiliary notation and terminology. Given a belief view $\mathbb{W}$, a *candidate unfounded set* $\mathbb{S}$ is a non-empty set of pairs of the form $\langle X, I\rangle$ such that $I \in \mathbb{W}$ and $X \cap I \neq \emptyset$. Furthermore, we write $\mathbb{W} - \mathbb{S}$ to denote the belief view:

$$\{\ \langle I, I\rangle \in \mathbb{W} \mid \langle X, I\rangle \notin \mathbb{S}\ \} \cup \{\ \langle I \setminus X, I\rangle \in \mathbb{W} \mid \langle I, I\rangle \in \mathbb{W} \text{ and } \langle X, I\rangle \in \mathbb{S}\ \}$$

Then, this belief view $\mathbb{W} - \mathbb{S}$ satisfies the following property:

Proposition 6. *Let* Π *be an epistemic program and* $\mathbb{W}$ *be a total belief view such that* $\mathbb{W} \models \Pi$*. If* $\mathbb{S}$ *is an unfounded set with respect to* Π *and* $\mathbb{W}$*, then* $\mathbb{W} - \mathbb{S} \models \Pi$*.* □

Proof. Suppose, for the sake of contradiction, that $\mathbb{S}$ is an unfounded set with respect to Π and $\mathbb{W}$ and $\mathbb{W} - \mathbb{S} \not\models \Pi$. Then, there is some rule $r \in \Pi$ such that $\mathbb{W} \models r$ and $\mathbb{W} - \mathbb{S} \models r$. In its turn, this implies that there is some pair $\langle H, T\rangle \in \mathbb{W} - \mathbb{S}$ such that $(\mathbb{W} - \mathbb{S}, H, T) \not\models r$ and, thus, that one of following conditions must hold:

1. $(\mathbb{W} - \mathbb{S}, H, T) \models \bigwedge Body(r)$ and $(\mathbb{W}', H, T) \not\models \bigvee Head(r)$, or
2. $(\mathbb{W}, I, I) \models \bigwedge Body(r)$ and $(\mathbb{W}, I, I) \not\models \bigvee Head(r)$.

Note that the latter is a contradiction with the fact that $\mathbb{W} \models r$. Furthermore, $(\mathbb{W} - \mathbb{S}, H, I) \models \bigwedge \mathit{Body}(r)$ implies $(\mathbb{W}, I) \models \bigwedge \mathit{Body}(r)$ and, since $\mathbb{W} \models r$, also that $(\mathbb{W}, I) \models \bigvee \mathit{Head}(r)$. Hence, $\mathit{Head}(r) \cap H = \emptyset$ and there is an atom $a \in \mathit{Head}(r)$ such that $a \in I \setminus H$. By construction, this implies that $\langle X, I \rangle \in \mathbb{S}$ with $X = I \setminus H$ and, thus, one of the following conditions holds:

3. $(\mathbb{W}, T) \not\models \bigwedge \mathit{Body}(r)$,
4. $\mathit{Body}^+_{ob}(r) \cap X \neq \emptyset$, or
5. $(\mathit{Head}(r) \setminus X) \cap T \neq \emptyset$, or
6. $\mathit{Body}^+_{sub}(r) \cap Y \neq \emptyset$.

Condition 3 cannot hold because it implies $(\mathbb{W}, T) \models \bigwedge \mathit{Body}(r)$. Furthermore, $X = I \setminus H$ implies that Condition 5 cannot hold either. The same happens for Condition 4, as we show next. Assume that there is some atom $b \in \mathit{Body}^+_{ob}(r) \cap X \neq \emptyset$. Then, we get $(\mathbb{W} - \mathbb{S}, H, T) \not\models b$ and, thus, also $(\mathbb{W} - \mathbb{S}, H, T) \not\models \bigwedge \mathit{Body}(r)$ which is a contradiction with Condition 1. Therefore, it must be that $\mathit{Body}^+_{sub}(r) \cap Y \neq \emptyset$ holds. Pick some atom $b \in \mathit{Body}^+_{sub}(r) \cap Y$. But then, there is some $\langle X', I' \rangle \in \mathbb{S}$ such that $b \in X'$ and $\langle H', I' \rangle \in \mathbb{W}$ with $X' = I' \setminus H'$. This implies that $(\mathbb{W}', H, I) \not\models \mathbf{K}\, b$ which is a contradiction with the fact that $(\mathbb{W}', H, I) \models \bigwedge \mathit{Body}(r)$. Consequently, $\mathbb{W} - \mathbb{S} \models \Pi$. □

This result is used next to prove that S5-EL satisfies foundedness:

Theorem 5. *Any S5-world-view of any program* Π *is founded.* □

Proof. Let $\mathbb{W}$ be some S5-world view of Π and suppose, for the sake of contradiction, that it is not founded. Then, there is a unfounded-set $\mathbb{S}$ for Π with respect to $\mathbb{W}$ such that every $\langle X, I \rangle \in \mathbb{S}$ satisfies $I \in \mathbb{W}$ and $X \cap I \neq \emptyset$. From Proposition 6, this implies that $\mathbb{W} - \mathbb{S} \models \Pi$ and it is easy to see that $\mathbb{W} - \mathbb{S} \prec \mathbb{W}$. This is a contradiction with the fact that $\mathbb{W}$ is an S5-world view of Π and, consequently, $\mathbb{W}$ must be founded. □

As an immediate consequence, since FAEEL world views are also S5-world views (Theorem 4) we directly conclude:

Theorem 6. *FAEEL satisfies foundedness.* □

Theorems 2 and 6 assert that any FAEEL-world view is a founded G94-world view. The natural question is whether the opposite also holds, that is, whether any founded G94-world view is also a FAEEL-world view. In Examples 6, 7 and 9 we did not find any counterexample, and this is in fact a general property that we will prove as our Main Theorem below. To this aim, we begin showing that the converse of Proposition 6 also holds. Given a belief view $\mathbb{W}$, let us define the following candidate unfounded set:

$$\mathbb{S}_{\mathbb{W}} \quad \stackrel{\text{def}}{=} \quad \{ \ \langle X, I \rangle \ \mid \ \langle H, I \rangle \in \mathbb{W} \text{ with } X = I \setminus H \text{ and } X \neq \emptyset \ \}$$

Proposition 7. *Let Π be an epistemic program and $\mathbb{W}$ be a non-total belief view such that $\mathbb{W} \models \Pi$. Then, $\mathbb{W} \models \Pi$ implies that $\mathbb{S}_{\mathbb{W}}$ is an unfounded set with respect to Π and $\mathbb{W}^t$.* □

Proof. Assume that $\mathbb{W} \models \Pi$ and suppose, for the sake of contradiction, that $\mathbb{S}_{\mathbb{W}}$ is not an unfounded set with respect to Π and $\mathbb{W}^t$. Then, there is some pair $\langle X, I \rangle \in \mathbb{S}$ and some rule $r \in \Pi$ with $\mathit{Head}(r) \cap X \neq \emptyset$ satisfying:

1. $(\mathbb{W}^t, I) \models \bigwedge \mathit{Body}(r)$, and
2. $\mathit{Body}^+_{ob}(r) \cap X = \emptyset$, and
3. $(\mathit{Head}(r) \setminus X) \cap I = \emptyset$.
4. $\mathit{Body}^+_{sub}(r) \cap Y = \emptyset$, and

with $Y = \bigcup \{ \ X' \mid \langle X', I' \rangle \in \mathbb{S} \ \}$. Furthermore, from $\mathbb{W} \models \Pi$, it immediately follows that $(\mathbb{W}, H, I) \models r$ for every rule $r \in \Pi$ and pair $\langle H, I \rangle \in \mathbb{W}$. Furthermore, since $\mathbb{W}$ is non-total, there is some $\langle H, I \rangle \in \mathbb{W}$ such that $H \subset I$ and, thus, $\langle X, I \rangle \in \mathbb{S}_{\mathbb{W}}$. Then, since $(\mathbb{W}, H, I) \models r$, one of the following conditions must hold:

5. $(\mathbb{W}, I) \not\models \bigwedge \mathit{Body}(r)$, or
6. $(\mathbb{W}, H, I) \models \bigvee \mathit{Head}(r)$, or
7. $(\mathbb{W}, H, I) \not\models \bigwedge \mathit{Body}(r)$ and $(\mathbb{W}, I) \models \bigvee \mathit{Head}(r)$.

Clearly, (5) is in contradiction with (1) and, thus, either (6) or (7) must hold. Note also that (3) implies that $\mathit{Head}(r) \cap H = \emptyset$ and, thus, (6) cannot hold either. Hence, (7) must hold and, thus, there is some literal $L \in \mathit{Body}^+(r)$

such that $(\mathbb{W}, H, I) \not\models L$. If $L \in Body_{ob}(r)$, then $L \notin H$. Besides, from Condition 1, it follows $L \in I$ and, thus, $L \in I \setminus H = X$ which is a contradiction with Condition 2. Otherwise, $L \in Body_{sub}(r)$ and there is some $\langle H', I' \rangle \in \mathbb{W}$ such that $a \notin H'$ with $L = \mathbf{K}\, a$. On the other hand, Condition 1 implies $(\mathbb{W}, I) \models \bigwedge Body(r)$ and, thus, $(\mathbb{W}, I) \models \mathbf{K}\, a$. In its turn, this implies that $a \in I'$ for every pair $\langle I', I' \rangle \in \mathbb{W}$. But then, by construction, there is some $\langle X', I' \rangle \in \mathbb{S}$ with $X' = I' \setminus H'$ and, thus, $a \in X' \subseteq Y$ which is a contradiction with Condition 4. Consequently, $\mathbb{S}_{\mathbb{W}}$ is an unfounded set with respect to Π and $\mathbb{W}$. $\square$

Next, we prove that a total belief model $\mathbb{W}$ of a program Π is founded if and only if $\mathbb{W}$ is an S5-world view.

Theorem 7. *Given any program* Π*, any belief view* $\mathbb{W}$ *is an S5-world view of* Π *iff* $\mathbb{W}$ *is a founded total belief model of* Π. $\square$

Proof. From Theorem 5, we obtain that any S5-world view is founded. Suppose now, for the sake of contradiction, that $\mathbb{W}$ is a founded total model of Π, but not an S5-world view. Then, there is some belief view $\mathbb{W}'$ such that $\mathbb{W}' \prec \mathbb{W}$ and $\mathbb{W} \models \Pi$ and, from Proposition 7, we obtain that $\mathbb{S}_{\mathbb{W}'}$ is an unfounded set with respect to Π and $\mathbb{W}$. This is a contradiction with the assumption that $\mathbb{W}$ is a founded total model of Π and, thus, it must be that $\mathbb{W}$ is an S5-world view of Π. $\square$

This last result reveals that the S5-world views selection can be somehow considered as the *semantic counterpart* of the foundedness (syntactic-dependent) condition. The S5-EL semantic characterisation of foundedness has the additional advantage of being applicable to any arbitrary theory and not just to epistemic programs.

Finally, we have now all the conditions to prove the Main Theorem:

Main Theorem. *Given any program* Π*, its equilibrium world views coincide with its founded G94-world views.* $\square$

Proof. From Theorem 4 it follows that $\mathbb{W}$ is a FAEEL-world view iff (i) $\mathbb{W}$ is a G94-world view and (ii) $\mathbb{W}$ is an S5-world view. On the other hand, Theorem 7 proves that (ii) is equivalent to require that $\mathbb{W}$ is a founded total belief model of Π. Finally, note that every G94-world view is also a total belief model of Π (see Proposition 8 in the next section) and, thus, we can rewrite the two conditions (i) and (ii) as: (i) $\mathbb{W}$ is a G94-world view and (ii) $\mathbb{W}$ is a founded G94-world view. $\square$

program	world views
$a \vee b$	$[\{a\},\{b\}]$
$a \vee b$ $a \leftarrow \mathbf{K}\, b$	$[\{a\},\{b\}]$
$a \vee b$ $a \leftarrow \neg\mathbf{K}\, b$	$[\{a\}]$
$a \vee b$ $c \leftarrow \neg\mathbf{K}\, b$	$[\{a,c\},\{b,c\}]$
$a \leftarrow \neg\mathbf{K}\, b$ $b \leftarrow \neg\mathbf{K}\, a$	$[\{a\}]$, $[\{b\}]$
$a \leftarrow$ $\neg\mathbf{K}\neg a$ $a \leftarrow \neg\mathbf{K}\, a$	$[\{a\}]$

program	G94/G11/FAEEL	K15/F15/S17
$a \leftarrow \neg\mathbf{K}\neg a$	$[\emptyset]$, $[\{a\}]$	$[\{a\}]$
$a \vee b$ $a \leftarrow \neg\mathbf{K}\neg b$	none	$[\{a\}]$
$a \vee b$ $a \leftarrow \mathbf{K}\neg b$	$[\{a\}]$, $[\{a\},\{b\}]$	$[\{a\},\{b\}]$
$a \leftarrow b$ $b \leftarrow \neg\mathbf{K}\neg a$	$[\emptyset]$, $[\{a,b\}]$	$[\{a,b\}]$
$a \leftarrow \neg\mathbf{K}\neg b$ $b \leftarrow \neg\mathbf{K}\neg a$	$[\emptyset]$, $[\{a\},\{b\}]$	$[\{a\},\{b\}]$

Figure 1: On the left, examples where G94, G11, K15, F15, S17 and FAEEL agree. On the right, examples where FAEEL/G91/G11 differ from K15/F15/S17.

8 Comparison to other approaches

To illustrate the effect of FAEEL when compared to other approaches in the literature, we begin providing a list of usual examples taken from Table 4 in [21]. The left table in Figure 1 contains programs where all semantics agree. The right table contains examples where the different semantics are divided into two groups: one in which G94, G11 and FAEEL coincide, and another in which the other three, K15, F15 and S17, coincide. Note that, in these programs, all subjective literals are negative and, by Theorem 3, we know that G94, G11 and FAEEL coincide for that syntactic class. The table in Figure 2 shows more cases where G11 and FAEEL coincide, but differ from one of the other semantics. Finally, Examples 1 and 7 in the paper can be used to differentiate between FAEEL and G11.

Lifting this comparison from mere examples to formal properties, we already saw that FAEEL is the only one that satisfies foundedness. Besides, Cabalar et al. [6] recently proposed other four properties for potential semantics of epistemic specifications. Among them, *epistemic splitting* is inspired by the well-known splitting theorem for standard logic programs [37]. Informally speaking, this property states that an epistemic logic program can be split if its *top* part only refers to the atoms of the *bottom* part through sub-

program	G94	G11/FAEEL	K15	F15/S17
$a \leftarrow \neg\mathbf{K}\neg b \wedge \neg b$ $b \leftarrow \neg\mathbf{K}\neg a \wedge \neg a$	$[\emptyset]$, $[\{a\},\{b\}]$			$[\{a\},\{b\}]$
$a \leftarrow \mathbf{K}\,a$	$[\emptyset]$, $[\{a\}]$	$[\emptyset]$		
$a \leftarrow \mathbf{K}\,a$ $a \leftarrow \neg\mathbf{K}\,a$	$[\{a\}]$	none		

Figure 2: Examples showing differences among several semantics. Examples 1 and 7 in the paper can be used to distinguish between FAEEL and G11.

jective literals. Then, a given semantics is said to satisfy epistemic splitting if it is possible to get its world views by first obtaining the world views of the bottom and then using the subjective literals in the top as "queries" on the bottom part previously obtained. For simplicity, we will not go here into the formal details of this property. The reader can check [6], where G94 is proved to satisfy epistemic splitting, and [19] where the same is proved for FAEEL.

We analyse here the other three properties introduced in [6]:

1. *supra-ASP* holds when, for any objective theory Γ, either: Γ has a unique world view $\mathbb{W} = \mathrm{SM}[\Gamma] \neq \emptyset$; or $\mathrm{SM}[\Gamma] = \emptyset$ and Γ has no world view.

2. *supra-S5* holds when every world view $\mathbb{W}$ of a theory Γ is also an epistemic model of Γ (that is, a model in the modal logic S5, $\mathbb{W} \models \Gamma$).

3. *subjective constraint monotonicity* holds when, for any theory Γ and any subjective constraint $\bot \leftarrow \varphi$, it follows that $\mathbb{W}$ is a world view of $\Gamma \cup \{\bot \leftarrow \varphi\}$ iff both $\mathbb{W}$ is a world view of Γ and $\mathbb{W}$ is not an S5-model of φ.

To prove that FAEEL satisfies these properties, we begin introducing three lemmata.

Lemma 4. *Let φ be a formula in which every atom is in the scope of the modal operator $\mathbf{K}$ and $\mathcal{I} = (\mathbb{W}, H, T)$ be some belief interpretation. Then, $\mathcal{I} \models \varphi$ iff $\mathbb{W} \models \varphi$.* □

Proof. In case $\varphi = \mathbf{K}\,\psi$, we can see that $\mathcal{I} \models \varphi$ iff $(\mathbb{W}, H', T') \models \psi$ for all $\langle H', T'\rangle \in \mathbb{W}$ iff $\mathbb{W} \models \varphi$. The rest of the proof follows by induction on the structure of φ. □

Lemma 5. *Let φ be a formula in which every atom is in the scope of the epistemic operator* $\mathbf{K}$ *and* $\mathcal{I} = (\mathbb{W}, H, T)$ *be some belief interpretation. Then, we get that* $\mathcal{I} \models \bot \leftarrow \varphi$ *iff* $\mathbb{W}^t \models \bot \leftarrow \varphi$. □

Proof. By definition, it follows that $\mathcal{I}$ is a model of $\bot \leftarrow \varphi$
iff $(\mathbb{W}, H', T') \models \bot \leftarrow \varphi$ for all $\langle H', T' \rangle \in \mathbb{W} \cup \{\langle H, T \rangle\}$
iff $(\mathbb{W}^t, H', T') \not\models \varphi$ for all $\langle H', T' \rangle \in \mathbb{W} \cup \{\langle H, T \rangle\}$
iff $\mathbb{W} \not\models \varphi$ (Lemma 4)
iff $\mathbb{W} \models \bot \leftarrow \varphi$
iff $\mathbb{W}^t \models \bot \leftarrow \varphi$ (Proposition 3). □

Lemma 6. *Let* Γ *be a theory and* φ *be a formula in which every atom is in the scope of the epistemic operator* $\mathbf{K}$ *and* $\mathcal{I} = (\mathbb{W}, H, T)$ *be some belief interpretation. Then,* $\mathcal{I}$ *is a equilibrium model of* $\Gamma \cup \{\bot \leftarrow \varphi\}$ *iff* $\mathcal{I}$ *is a equilibrium model of* Γ *and* $\mathbb{W}^t \models \bot \leftarrow \varphi$. □

Proof. Assume first that $\mathcal{I}$ is a equilibrium model of $\Gamma \cup \{\bot \leftarrow \varphi\}$. Then, $\mathcal{I}$ is a model of Γ and a model of $\bot \leftarrow \varphi$. From Lemma 5, the latter implies that $\mathbb{W}^t \models \bot \leftarrow \varphi$. Suppose, for the sake of contradiction, that $\mathcal{I}$ is a equilibrium model of Γ and, thus, that there is some model $\mathcal{I}' = (\mathbb{W}', H', T')$ of Γ such that $\mathcal{I}' \prec \mathcal{I}$. Then, $\mathcal{I}' \prec \mathcal{I}$ implies that $(\mathbb{W}')^t = \mathbb{W}$ and, thus, from Lemma 5 and the fact that $\mathbb{W}^t \models \bot \leftarrow \varphi$, we obtain that $\mathcal{J} \models \bot \leftarrow \varphi$. This contradicts the fact that $\mathcal{I}$ is an equilibrium model of $\Gamma \cup \{\bot \leftarrow \varphi\}$.

The other way around, assume that $\mathcal{I}$ is an equilibrium model of Γ and $\mathbb{W}^t \models \bot \leftarrow \varphi$. From Lemma 5, $\mathcal{I}$ is a model of $\bot \leftarrow \varphi$ and, thus, an equilibrium model of $\Gamma \cup \{\bot \leftarrow \varphi\}$. □

Proposition 8. *FAEEL satisfies supra-ASP, supra-S5 and subjective constraint monotonicity.* □

Proof. For *supra-ASP*, note that since we are dealing with propositional theories there is no occurrence of $\mathbf{K}$ and, thus, we can apply Theorem 3 to see that FAEEL and G94-world views coincide. Then, just note that, since there is no occurrence of $\mathbf{K}$, it follows that $\Gamma^{\mathbb{W}} = \Gamma$ and, thus $\mathbb{W}$ is an autoepistemic world view of Γ iff $\mathbb{W}$ is a G94-world view of Γ iff $\mathbb{W} = \mathrm{SM}[\Gamma^{\mathbb{W}}] = \mathrm{SM}[\Gamma]$.

For *supra-S5*, from Theorem 4, it follows that every FAEEL-world view $\mathbb{W}$ is also an S5-world view and it is easy to check that every S5-world view is also an epistemic model.

	G94	G11	L15	K15	S17	FAEEL
Supra-ASP	✓	✓	✓	✓	✓	✓
Supra-S5	✓	✓	✓	✓	✓	✓
Subjective constraint monotonicity	✓	✓				✓
Splitting	✓					✓
Foundedness						✓

Table 1: Summary of properties in different semantics [19]. By G94 we refer to the semantics of [23, 59, 62] since all of them agree when theories are epistemic programs. G11, F15, K15, and S17 correspond to the semantics in [24, 58, 34, 56], respectively.

For *subjective constraint monotonicity*, note that $\mathbb{W}$ is a FAEEL-world view of $\Gamma \cup \{\bot \leftarrow \varphi\}$ iff the following equality holds:

$$\mathbb{W} \quad = \quad \{\ T \mid (\mathbb{W}, T) \in \mathrm{EQB}[\Gamma \cup \{\bot \leftarrow \varphi\}]\ \}$$

Furthermore, from Lemma 6 and the fact that every atom in φ is in the scope of the modal operator $\mathbf{K}$, we conclude $(\mathbb{W}, T) \in \mathrm{EQB}[\Gamma \cup \{\bot \leftarrow \varphi\}]$ iff $(\mathbb{W}, T) \in \mathrm{EQB}[\Gamma]$ and $\mathbb{W} \models \bot \leftarrow \varphi$. Hence, $\mathbb{W}$ is a FAEEL-world view of $\Gamma \cup \{\bot \leftarrow \varphi\}$ iff the following equality holds:

$$\mathbb{W} \quad = \quad \{\ T \mid (\mathbb{W}, T) \in \mathrm{EQB}[\Gamma]\ \}$$

and $\mathbb{W} \models \bot \leftarrow \varphi$. □

All semantics discussed in this paper satisfy the above first two properties (supra-ASP and supra-S5), but most of them fail for subjective constraint monotonicity, as first discussed in [33]. In fact, a variation of Example 7 can be used to show that K15, F15 and S17 do not satisfy this property.

Example 10 (Example 7 continued)**.** *Suppose we remove the constraint (last rule) from Π_3 getting the program $\Pi_4 = \{a \vee b\,,\ c \leftarrow \mathbf{K}\, a\}$. All semantics, including G94 and FAEEL, agree that Π_4 has a unique world view $[\{a\}, \{b\}]$. Suppose we add now a subjective constraint $\Pi_5 = \Pi_4 \cup \{\bot \leftarrow \neg\mathbf{K}\, c\}$. This addition leaves G94, G11 and FAEEL without world views (due to subjective constraint monotonicity), but not for K15, F15 and S17, which provide a* new *world view $[\{a, c\}]$, not present before the addition of the constraint.* □

Table 1 summarises all the properties we discussed here and whether a semantics satisfies it or not. FAEEL is unique in guaranteeing foundedness while also satisfying epistemic splitting. This places FAEEL as a firm candidate for the semantics of epistemic logic programs. Furthermore, we have seen (Theorem 3) that, for epistemic tight programs (those not containing cycles involving positive epistemic literals), both G94 and FAEEL coincide. Recall that what made G94 inconvenient in the first place was the presence of self-supported world views due to positive cycles. Hence, this shows that FAEEL "fixes" the problem with self-supported world views present in G94 without introducing further variations that are unrelated to this problem.

9 Conclusions

In order to characterise self-supported world-views, already present in Gelfond's 1991 semantics [23] (G94), we have extended the definition of unfounded sets from standard logic programs to epistemic specifications. As a result, we proposed the *foundedness* property for epistemic semantics, which is not satisfied by other approaches in the literature. Our main contribution has been the definition of a new semantics, based on the so-called *Founded Autoepistemic Equilibrium Logic* (FAEEL), that satisfies foundedness. This semantics actually covers the syntax of any arbitrary modal theory and is a combination of Equilibrium Logic and Autoepistemic Logic. As a main result, we were able to prove that, for the syntax of epistemic specifications, FAEEL world views coincide with the set of G94 world views that are founded. We showed how this semantics behaves on a set of common examples in the literature and proved that it satisfies other four basic properties: all world views are S5 models (*supra-S5*); standard programs have (at most) a unique world view containing all the stable models (*supra-ASP*); subjective constraints just remove world views (*subjective constraint monotonicity*); and world views of programs can be computed modularly for those programs that can be split (*epistemic splitting*).

Our immediate future work is focused on the implementation of an efficient tool to compute FAEEL world views by including a foundedness check on top of some G94 solver, like for instance the one presented in [11]. The latter is an epistemic extension of the popular ASP solver `clingo` [22], so it can reuse many existing encodings and adapt them to solve problems that deal with the agent's knowledge and beliefs or with conformant versions of existing planning problems.

Acknowledgements. We are grateful to David Pearce and the anonymous reviewers from the journal and, previously, from the *15th International Conference on Logic Programming and Nonmonotonic Reasoning* (where a preliminary version of this work was presented), for their comments and suggestions that have helped us to improve the paper. We are also thankful to the organisers of LPNMR'19 for their support in preparing this extended version. This research was partially supported by MINECO, Spain, grant TIC2017-84453-P, Xunta de Galicia, Spain (GPC ED431B 2019/03). The second author was funded by the Centre International de Mathématiques et d'Informatique de Toulouse (CIMI) through contract ANR-11-LABEX-0040-CIMI within the program ANR-11-IDEX-0002-02 and the Alexander von Humboldt Foundation.

References

[1] F. Aguado, P. Cabalar, M. Diéguez, G. Pérez, and C. Vidal. Temporal equilibrium logic: a survey. *Journal of Applied Non-Classical Logics*, 23(1-2):2–24, 2013.

[2] Felicidad Aguado, Pedro Cabalar, Jorge Fandinno, David Pearce, Gilberto Pérez, and Concepción Vidal. Forgetting auxiliary atoms in forks. *Artificial Intelligence*, 275:575–601, 2019.

[3] L. Aiello and F. Massacci. Verifying security protocols as planning in logic programming. *ACM Transactions on Computational Logic*, 2(4):542–580, 2001.

[4] Giovanni Amendola, Thomas Eiter, Michael Fink, Nicola Leone, and João Moura. Semi-equilibrium models for paracoherent answer set programs. *Artificial Intelligence*, 234:219–271, 2016.

[5] Philippe Balbiani, Martín Diéguez, and Luis Fariñas del Cerro. Setting the basis for here and there modal logics. *FLAP*, 6(6):1475–1500, 2019.

[6] P. Cabalar, J. Fandinno, and L. Fariñas del Cerro. Splitting epistemic logic programs. *Theory and Practice of Logic Programming*, page 121, 2020.

[7] P. Cabalar, J. Fandinno, T. Schaub, and S. Schellhorn. Gelfond-zhang aggregates as propositional formulas. In M. Balduccini and T. Janhunen, editors, *Proceedings of the Fourteenth International Conference*

on Logic Programming and Nonmonotonic Reasoning (LPNMR'17), volume 10377 of *Lecture Notes in Artificial Intelligence*, pages 117–131. Springer-Verlag, 2017.

[8] P. Cabalar, J. Fandinno, T. Schaub, and S. Schellhorn. Lower bound founded logic of here-and-there. In F. Calimeri, N. Leone, and M. Manna, editors, *Proceedings of the Sixteenth European Conference on Logics in Artificial Intelligence (JELIA'19)*, volume 11468 of *Lecture Notes in Computer Science*, pages 509–525. Springer-Verlag, 2019.

[9] P. Cabalar, R. Kaminski, M. Ostrowski, and T. Schaub. An ASP semantics for default reasoning with constraints. In R. Kambhampati, editor, *Proceedings of the Twenty-fifth International Joint Conference on Artificial Intelligence (IJCAI'16)*, pages 1015–1021. IJCAI/AAAI Press, 2016.

[10] Pedro Cabalar, Jorge Fandinno, Luis Fariñas del Cerro, and David Pearce. Functional ASP with intensional sets: Application to Gelfond-Zhang aggregates. *Theory Pract. Log. Program.*, 18(3-4):390–405, 2018.

[11] Pedro Cabalar, Jorge Fandinno, Javier Garea, Javier Romero, and Torsten Schaub. A solver for epistemic specifications under gelfond 1991 semantics. In *Proceedings of the First Declarative Problem Solving Workshop (DPSW 2020), Santiago de Compostela, Spain, August 30th*, 2020.

[12] Pedro Cabalar, Sergei P. Odintsov, David Pearce, and Agustín Valverde. Partial equilibrium logic. *Ann. Math. Artificial Intelligence*, 50(3-4):305–331, 2007.

[13] Ernest Davis and Leora Morgenstern. Epistemic logic and its applications. In *Tutorial Notes, International Joint Conference on Artificial Intelligence (IJCAI'93)*, 1993.

[14] Jos de Bruijn, David Pearce, Axel Polleres, and Agustín Valverde. A semantical framework for hybrid knowledge bases. *Knowledge Information Systems*, 25(1):81–104, 2010.

[15] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*. Springer, 1st edition, 2007.

[16] T. Eiter, N. Leone, and D. Pearce. Assumption sets for extended logic programs. In J. Gerbrandy, M. Marx, M. de Rijke, and Y. Venema, editors, *Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday*, 1999.

[17] J. Engelfriet. Minimal temporal epistemic logic. *Notre Dame Journal of Formal Logic*, 37:233–259, 1996.

[18] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning About Knowledge*. MIT Press, Cambridge, MA, USA, 2003. Revised paperback edition published in 2003 from the 1995 edition.

[19] J. Fandinno. Founded (auto)epistemic equilibrium logic satisfies epistemic splitting. *Theory and Practice of Logic Programming*, 19(5-6):671–687, 2019.

[20] Jorge Fandinno and Luis Fariñas del Cerro. Constructive logic covers argumentation and logic programming. In Michael Thielscher, Francesca Toni, and Frank Wolter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the Sixteenth International Conference, KR 2018, Tempe, Arizona, 30 October - 2 November 2018*, pages 128–137. AAAI Press, 2018.

[21] L. Fariñas del Cerro, A. Herzig, and E. Iraz Su. Epistemic equilibrium logic. In Q. Yang and M. Wooldridge, editors, *Proceedings of the Twenty-fourth International Joint Conference on Artificial Intelligence (IJCAI'15)*, pages 2964–2970. AAAI Press, 2015.

[22] M. Gebser, R. Kaminski, B. Kaufmann, M. Ostrowski, T. Schaub, and P. Wanko. Theory solving made easy with clingo 5. In M. Carro and A. King, editors, *Technical Communications of the Thirty-second International Conference on Logic Programming (ICLP'16)*, volume 52 of *OpenAccess Series in Informatics (OASIcs)*, pages 2:1–2:15. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2016.

[23] M. Gelfond. Strong introspection. In T. Dean and K. McKeown, editors, *Proceedings of the Ninth National Conference on Artificial Intelligence (AAAI'91)*, pages 386–391. AAAI Press / The MIT Press, 1991.

[24] M. Gelfond. New semantics for epistemic specifications. In J. Delgrande and W. Faber, editors, *Proceedings of the Eleventh International*

Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'11), volume 6645 of *Lecture Notes in Artificial Intelligence*, pages 260–265. Springer-Verlag, 2011.

[25] M. Gelfond and V. Lifschitz. The stable model semantics for logic programming. In R. Kowalski and K. Bowen, editors, *Proceedings of the Fifth International Conference and Symposium of Logic Programming (ICLP'88)*, pages 1070–1080. MIT Press, 1988.

[26] M. Gelfond and H. Przymusinska. Reasoning on open domains. In Moniz Pereira and Nerode [45], pages 397–413.

[27] M. Gelfond and H. Przymusinska. Reasoning on open domains. In Moniz Pereira and Nerode [45], pages 397–413.

[28] Michael Gelfond. On stratified autoepistemic theories. In Kenneth D. Forbus and Howard E. Shrobe, editors, *Proceedings of the 6th National Conference on Artificial Intelligence. Seattle, WA, USA, July 1987*, pages 207–211. Morgan Kaufmann, 1987.

[29] Michael Gelfond. Logic programming and reasoning with incomplete information. *Annals of Mathematics and Artificial Intelligence*, 12(1-2):89–116, 1994.

[30] Amelia Harrison, Vladimir Lifschitz, David Pearce, and Agustín Valverde. Infinitary equilibrium logic and strongly equivalent logic programs. *Artificial Intelligence*, 246:22–33, 2017.

[31] A. Heyting. Die formalen Regeln der intuitionistischen Logik. *Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse*, pages 42–56, 1930.

[32] J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca, NY, 1962.

[33] P. Kahl and A. Leclerc. Epistemic logic programs with world view constraints. In A. Dal Palù, P. Tarau, N. Saeedloei, and P. Fodor, editors, *Technical Communications of the Thirty-fourth International Conference on Logic Programming (ICLP'18)*, volume 64 of *OASICS*, pages 1:1–1:17. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018.

[34] P. Kahl, R. Watson, E. Balai, M. Gelfond, and Y. Zhang. The language of epistemic specifications (refined) including a prototype solver. *Journal of Logic and Computation*, 2015.

[35] K. Konolige. On the relation between default and autoepistemic logic. *Artificial Intelligence*, 35(2):343–382, 1988.

[36] N. Leone, P. Rullo, and F. Scarcello. Disjunctive stable models: Unfounded sets, fixpoint semantics, and computation. *Information and Computation*, 135(2):69–112, 1997.

[37] V. Lifschitz and H. Turner. Splitting a logic program. In *Proceedings of the Eleventh International Conference on Logic Programming*, pages 23–37. MIT Press, 1994.

[38] Vladimir Lifschitz. Minimal belief and negation as failure. *Artificial Intelligence*, 70(1-2):53–72, 1994.

[39] Vladimir Lifschitz and Grigori Schwarz. Extended logic programs as autoepistemic theories. In Luís Moniz Pereira and Anil Nerode, editors, *, Proceedings of the Second International Workshop on Logic Programming and Non-monotonic Reasoning, Lisbon, Portugal, June 1993*, pages 101–114. MIT Press, 1993.

[40] Fangzhen Lin and Yoav Shoham. A logic of knowledge and justified assumptions. *Artificial Intelligence*, 57(2-3):271–289, 1992.

[41] V. Marek and M. Truszczyński. Relating autoepistemic and default logics. In R. Brachman, H. Levesque, and R. Reiter, editors, *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning (KR'89)*, pages 276–288. Morgan Kaufmann Publishers, 1989.

[42] V. Marek and M. Truszczyński. Stable models and an alternative logic programming paradigm. In K. Apt, V. Marek, M. Truszczyński, and D. Warren, editors, *The Logic Programming Paradigm: a 25-Year Perspective*, pages 375–398. Springer-Verlag, 1999.

[43] Victor W. Marek and Mirosław Truszczyński. *Nonmonotonic logic - context-dependent reasoning*. Artificial intelligence. Springer, 1993.

[44] Drew V. McDermott and Jon Doyle. Non-monotonic logic I. *Artificial Intelligence*, 13(1-2):41–72, 1980.

[45] L. Moniz Pereira and A. Nerode, editors. *Logic Programming and Non-monotonic Reasoning, Proceedings of the Second International Workshop, Lisbon, Portugal, June 1993*. MIT Press, 1993.

[46] R. Moore. Possible world semantics for autoepistemic logic. In *Proceedings of the AAAI-Workshop on Nonmonotonic Reasoning*, pages 344–354, 1984.

[47] R. Moore. Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25:75–94, 1985.

[48] I. Niemelä. Constructive tightly grounded autoepistemic reasoning. In J. Mylopoulos and R. Reiter, editors, *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI'91)*, pages 399–405. Morgan Kaufmann, 1991.

[49] I. Niemelä. Logic programs with stable model semantics as a constraint programming paradigm. *Annals of Mathematics and Artificial Intelligence*, 25(3-4):241–273, 1999.

[50] D. Pearce. A new logical characterisation of stable models and answer sets. In J. Dix, L. Pereira, and T. Przymusinski, editors, *Proceedings of the Sixth International Workshop on Non-Monotonic Extensions of Logic Programming (NMELP'96)*, volume 1216 of *Lecture Notes in Computer Science*, pages 57–70. Springer-Verlag, 1997.

[51] D. Pearce. Equilibrium logic. *Annals of Mathematics and Artificial Intelligence*, 47(1-2):3–41, 2006.

[52] Riccardo Pucella. Knowledge and security. *CoRR*, abs/1305.0876, 2013.

[53] Steven Schockaert, Jeroen Janssen, and Dirk Vermeir. Fuzzy equilibrium logic: Declarative problem solving in continuous domains. *ACM Transactions on Computational Logic*, 13(4):33:1–33:39, 2012.

[54] G. Schwarz. Autoepistemic logic of knowledge. In A. Nerode, V. Marek, and V. Subrahmanian, editors, *Proceedings of the First International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'91)*, pages 260–274. The MIT Press, 1991.

[55] G. Schwarz. Minimal model semantics for nonmonotonic modal logics. In *Proceedings of the Seventh Annual Symposium on Logic in Computer Science (LICS'92)*, pages 34–43. IEEE Computer Society, 1992.

[56] Y. Shen and T. Eiter. Evaluating epistemic negation in answer set programming. *Artificial Intelligence*, 237:115–135, 2016.

[57] T. Son, T. Le, P. Kahl, and A. Leclerc. On computing world views of epistemic logic programs. In C. Sierra, editor, *Proceedings of the Twenty-sixth International Joint Conference on Artificial Intelligence (IJCAI'17)*, pages 1269–1275. IJCAI/AAAI Press, 2017.

[58] Ezgi Iraz Su, Luis Fariñas del Cerro, and Andreas Herzig. Autoepistemic equilibrium logic and epistemic specifications. *Artificial Intelligence*, 282:103249, 2020.

[59] Mirosław Truszczyński. Revisiting epistemic specifications. In M. Balduccini and T. Son, editors, *Logic Programming, Knowledge Representation, and Nonmonotonic Reasoning - Essays Dedicated to Michael Gelfond on the Occasion of His Sixty-fifth Birthday*, volume 6565 of *Lecture Notes in Computer Science*, pages 315–333. Springer, 2011.

[60] Hans van Ditmarsch, Joseph Y. Halpern, Wiebe van der Hoek, and Barteld Kooi, editors. *Handbook of Epistemic Logic*, pages 32–33. Applied Mathematics Series 55. London: College Publications, Washington, DC, 1964.

[61] A. Van Gelder, K. Ross, and J. Schlipf. The well-founded semantics for general logic programs. *Journal of the ACM*, 38(3):620–650, 1991.

[62] K. Wang and Y. Zhang. Nested epistemic logic programs. In C. Baral, G. Greco, N. Leone, and G. Terracina, editors, *Proceedings of the Eighth International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'05)*, volume 3662 of *Lecture Notes in Artificial Intelligence*, pages 279–290. Springer-Verlag, 2005.