

# *Founded (Auto)Epistemic Equilibrium Logic Satisfies Epistemic Splitting\**

JORGE FANDINNO

IRIT, University of Toulouse, CNRS, France

(e-mail: [jorge.fandinno@irit.fr](mailto:jorge.fandinno@irit.fr))

Universität Potsdam, Germany

(e-mail: [fandinno@uni-potsdam.de](mailto:fandinno@uni-potsdam.de))

submitted 23 July 2019; accepted 31 July 2019

---

## Abstract

In a recent line of research, two familiar concepts from logic programming semantics (unfounded sets and splitting) were extrapolated to the case of epistemic logic programs. The property of *epistemic splitting* provides a natural and modular way to understand programs without epistemic cycles but, surprisingly, was only fulfilled by Gelfond’s original semantics (G91), among the many proposals in the literature. On the other hand, G91 may suffer from a kind of self-supported, unfounded derivations when epistemic cycles come into play. Recently, the absence of these derivations was also formalised as a property of epistemic semantics called *foundedness*. Moreover, a first semantics proved to satisfy foundedness was also proposed, the so-called *Founded Autoepistemic Equilibrium Logic* (FAEEL). In this paper, we prove that FAEEL also satisfies the epistemic splitting property something that, together with foundedness, was not fulfilled by any other approach up to date. To prove this result, we provide an alternative characterisation of FAEEL as a combination of G91 with a simpler logic we called *Founded Epistemic Equilibrium Logic* (FEEL), which is somehow an extrapolation of the stable model semantics to the modal logic S5.

**KEYWORDS:** Answer Set Programming, Epistemic Specifications, Epistemic Logic Programs

---

## 1 Introduction

The language of *epistemic specifications* (Gelfond 1991) or *epistemic logic programs* extends disjunctive logic programs, under the *stable model* (Gelfond and Lifschitz 1988) semantics, with modal constructs called *subjective literals*. These constructs allow to check whether a regular or *objective 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  $a$  should be derived whenever we cannot prove that all the stable models contain  $b$ . The definition of a “satisfactory” semantics for epistemic specifications has proven to be a non-trivial enterprise with a long list of alternative semantics (Gelfond 1991; Wang and Zhang 2005; Truszczyński 2011; Gelfond 2011; Fariñas del Cerro et al. 2015;

\* This work has been partially supported by the Centre International de Mathématiques et d’Informatique de Toulouse (CIMI) through contract ANR-11-LABEX-0040-CIMI within the programme ANR-11-IDEX-0002-02 and the Alexander von Humboldt Foundation.

Kahl et al. 2015; Shen and Eiter 2017; Cabalar et al. 2019a). 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:

$$b \leftarrow \neg \mathbf{K} a \quad (2)$$

plus the above rule (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 (1991) (G91) considered different alternative *world views* or sets of stable models. In the case of program (1)-(2), G91 yields two alternative world views<sup>1</sup>,  $[\{a\}]$  and  $[\{b\}]$ , each one containing a single stable model, and this is also the behaviour obtained in the remaining approaches developed later on. As noted by (Truszczyński 2011), the feature that made G91 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:

$$a \leftarrow \mathbf{K} a \quad (3)$$

whose world views under G91 are  $[\emptyset]$  and  $[\{a\}]$ . The latter is considered as counter-intuitive by all authors<sup>2</sup> 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. Recently, Cabalar et al. (2019a) proposed to characterise these unintended world views by extending the notion of *unfounded sets* (Gelder et al. 1991) from standard disjunctive logic programs (Leone et al. 1997) to the case of epistemic logic programs. In that work, the authors also provided a new semantics, called *Founded Autoepistemic Equilibrium Logic* (FAEEL), that fulfills that requirement. In fact, FAEEL-world views are precisely those G91-world views that are founded, that is, those that do not admit any unfounded set.

On the other hand, it is obvious that programs without epistemic cycles (i.e. cycles involving epistemic literals) cannot have self-supported derivations. In this sense, one could expect that proposals that tried to get rid of G91 self-supported derivations coincided with the latter, at least, for epistemically acyclic programs. However, (Leclerc and Kahl 2018) have recently pointed out that this is not the case: for instance, while in G91, (purely) epistemic constraints always remove world views, this does not hold in other semantics. Watson (2000) and Cabalar et al. (2018; 2019b) went a step farther defining a property called *epistemic splitting* which, not only defines an intuitive behaviour for stratified epistemic specifications, but also extends the splitting theorem, well-known for autoepistemic logic (Gelfond and Przymusinska 1992) and standard logic programs (Lifschitz and Turner 1994), to the case of epistemic logic programs. For instance, if we consider a program consisting of rules (1)-(2) plus

$$c \leftarrow \mathbf{K} a \quad (4)$$

we may expect to obtain the world views  $[\{a, c\}]$  and  $[\{b\}]$  resulting from adding the atom  $c$  only to the belief sets of the world view that satisfies  $\mathbf{K} a$ . This property is known to be satisfied by the G91 semantics, but surprisingly not for those that tried to correct its self-supported problem (Cabalar et al. 2018; Cabalar et al. 2019b).

The major contribution of this paper is the proof that FAEEL satisfies the epistemic splitting property as defined in (Cabalar et al. 2018; Cabalar et al. 2019b). Joining this

<sup>1</sup> For the sake of readability, sets of propositional interpretations are embraced with  $[\ ]$  rather than  $\{ \}$ .

<sup>2</sup> This includes Gelfond himself, who proposed a new variant in (Gelfond 2011) motivated by this same example and further modified this variant later on in (Kahl et al. 2015).

result with the already known fact that this semantics also satisfies the foundedness property shows that FAEEL is a solid candidate to serve as a semantics of epistemic logic programs. A second contribution of this paper is the introduction of a logic that we call *Founded Epistemic Equilibrium Logic* (FEEL) and which can be intuitively seen as the combination of the Equilibrium Logic with the modal logic S5. For the sake of comparative, FAEEL corresponds to the combination of the Equilibrium Logic with the Moore’s Autoepistemic Logic (AEL; Moore 1985). In this sense, FEEL is the combination of a non-monotonic logic with a monotonic one, while FAEEL is the combination of two non-monotonic logics, a fact that makes FEEL much easier to study. This brings us to the third contribution of the paper: FAEEL world views can be precisely characterised as those G91 world views that are at the same time FEEL world views. This allows us to study FAEEL properties by studying them independently in FEEL and G91 and then combining their results. This is precisely the methodology used in proving the epistemic splitting theorem for FAEEL.

The rest of the paper is organised as follows. Section 2 revisits the background knowledge about equilibrium logic, epistemic specifications, the epistemic splitting property and FAEEL necessary for the rest of the paper. Section 3 introduces FEEL and studies the relation between this logic and FAEEL. In Section 4, we study the epistemic splitting property in FEEL and FAEEL and, in Section 5, we discuss other existent approaches to epistemic logic programs. Finally, Section 6 concludes the paper.

## 2 Background

We start by recalling the basic definitions needed for the rest of the paper. Given a set of atoms  $At$ , an (epistemic) formula is defined according to the following grammar:

$$\varphi ::= \perp \mid a \mid \varphi_1 \wedge \varphi_2 \mid \varphi_1 \vee \varphi_2 \mid \varphi_1 \rightarrow \varphi_2 \mid \mathbf{K} \varphi \quad \text{for any atom } a \in At.$$

In our context, the epistemic reading of  $\mathbf{K}\psi$  is that “ $\psi$  is one of the agent’s beliefs.” Thus, a formula  $\varphi$  is said to be *subjective* if all its atom occurrences (having at least one) are in the scope of  $\mathbf{K}$ . Analogously,  $\varphi$  is said to be *objective* if  $\mathbf{K}$  does not occur in  $\varphi$ . 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. Given a formula  $\varphi$ , by  $Atoms(\varphi)$  we denote the set of all atoms occurring in  $\varphi$ . For instance,  $Atoms(\neg a \vee \mathbf{K}b) = \{a, b\}$ . As usual we define the following derived operators:  $\varphi \leftrightarrow \psi \stackrel{\text{def}}{=} (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$ ,  $(\varphi \leftarrow \psi) \stackrel{\text{def}}{=} (\psi \rightarrow \varphi)$ ,  $\neg\varphi \stackrel{\text{def}}{=} (\varphi \rightarrow \perp)$  and  $\top \stackrel{\text{def}}{=} \neg\perp$ . An (epistemic) theory is a (possibly infinite) set of formulas as defined above and an *objective theory* is a theory whose formulas are objective. We write  $Atoms(\varphi)$  to represent the set of atoms occurring in any formula  $\varphi$  and  $Atoms(\Gamma)$  to represent the set of atoms occurring in any theory  $\Gamma$ . Recall that (Gelfond 1991) included a second subjective operator  $\mathbf{M}$  such that  $\mathbf{M}l$  is read as “the agent believes that  $l$  is possible.” In this paper, we assume here that  $\mathbf{M}\varphi$  is just an abbreviation<sup>3</sup> for  $\neg\mathbf{K}\neg\varphi$ .

### 2.1 Equilibrium Logic and the Stable Models Semantics

A *propositional interpretation*  $T$  is a set of atoms  $T \subseteq At$ . We write  $T \models \varphi$  to represent that  $T$  classically satisfies formula  $\varphi$ . An *HT-interpretation* is a pair  $\langle H, T \rangle$  (respectively

<sup>3</sup> Several interpretations of  $\mathbf{M}$  are possible in the logics considered in this paper depending on the level of foundedness that it is expected to satisfy. We limit ourselves here to the simplest interpretation of  $\mathbf{M}$ , leaving other interpretations for a more detailed discussion in the future.

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  $\varphi$  under the recursive conditions:

- $\langle H, T \rangle \not\models \perp$
- $\langle H, T \rangle \models p$  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  $\langle H, T \rangle \models \psi$
- $\langle H, T \rangle \models \varphi \rightarrow \psi$  iff both (i)  $T \models \varphi \rightarrow \psi$  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  $\Gamma$ , 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 ‘ $\models$ ’ indistinctly. Interpretation  $\langle T, T \rangle = T$  is a *stable (or equilibrium) model* of a theory  $\Gamma$  iff  $T \models \Gamma$  and there is no  $H \subset T$  such that  $\langle H, T \rangle \models \Gamma$ . We write  $\text{SM}[\Gamma]$  to stand for the set of all stable models of  $\Gamma$ .

## 2.2 G91 semantics for epistemic theories

To represent the agent’s beliefs, we will use a set  $\mathbb{W}$  of propositional interpretations. We call *belief set* to each element  $I \in \mathbb{W}$  and *belief view* the whole set  $\mathbb{W}$ . 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:

- $(\mathbb{W}, I) \not\models \perp$ ,
- $(\mathbb{W}, I) \models a$  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  $(\mathbb{W}, I) \models \psi_2$ ,
- $(\mathbb{W}, I) \models \psi_1 \rightarrow \psi_2$  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,  $\varphi \rightarrow \psi$  is equivalent to  $\neg\varphi \vee \psi$  in this context. A belief interpretation  $(\mathbb{W}, I)$  is a *belief model* of  $\Gamma$  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  $\Gamma$ , 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  $\Gamma_1 = \{\neg\mathbf{K}b \rightarrow a\}$  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  $\mathbf{K}b$ ,  $\{\{b\}\}$ ,  $\{\{a, b\}\}$ , and  $\{\{b\}, \{a, b\}\}$  and the rest of cases must force  $a$  true, so we also get  $\{\{a\}\}$  and  $\{\{a\}, \{a, b\}\}$ . In other words,  $\Gamma_1$  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 interested only in some epistemic models (that as usual 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 that extend the one given by (Truszczyński 2011) by stating an explicit signature in which it is applied. The explicit signature will be useful later on to define the epistemic splitting property.

**Definition 1** (Subjective reduct). *The subjective reduct of a theory  $\Gamma$  with respect to a set of belief views  $\mathbb{W}$  and a signature  $U \subseteq \text{At}$ , also written  $\Gamma_U^{\mathbb{W}}$ , is obtained by replacing each maximal subjective formula of the form  $\mathbf{K}\varphi$  with  $\text{Atoms}(\varphi) \subseteq U$  by  $\top$  if  $\mathbb{W} \models \mathbf{K}\varphi$ ; or by  $\perp$  otherwise. When  $U = \text{At}$  we just write  $\Gamma^{\mathbb{W}}$ .  $\square$*

Finally, we impose a fixpoint condition where each belief set  $I \in \mathbb{W}$  is required to be a stable model of the reduct, obtaining thus the G91 semantics.

**Definition 2** (G91 world view). *A belief view  $\mathbb{W}$  is called a G91-world view of  $\Gamma$  if and only if it satisfies  $\mathbb{W} = \text{SM}[\Gamma^{\mathbb{W}}]$ .  $\square$*

**Example 2** (Example 1 continued). *Take any  $\mathbb{W}$  such that  $\mathbb{W} \models \mathbf{K}b$ . Then,  $\Gamma_1^{\mathbb{W}} = \{\perp \rightarrow a\}$  with  $\text{SM}[\Gamma_1^{\mathbb{W}}] = [\emptyset]$ . The empty set does not satisfy  $\mathbf{K}b$  so  $\mathbb{W}$  cannot be a G91-world view of  $\Gamma_1$ . If  $\mathbb{W} \not\models \mathbf{K}b$  instead, we get  $\Gamma_1^{\mathbb{W}} = \{\top \rightarrow a\}$ , whose unique stable model is  $\{a\}$ . As a result,  $\mathbb{W} = [\{a\}]$  is the unique G91-world view.  $\square$*

**Example 3** (Example 2 continued). *Let now  $\Gamma_2 = \{\neg\mathbf{K}b \rightarrow a, \neg\mathbf{K}a \rightarrow b\}$  corresponding to rules (1)-(2). Take any  $\mathbb{W}$  such that  $\mathbb{W} \models \neg\mathbf{K}a \wedge \mathbf{K}b$ . Then,  $\Gamma_1^{\mathbb{W}} = \{\perp \rightarrow a, \top \rightarrow b\}$  and we have that  $\text{SM}[\Gamma_1^{\mathbb{W}}] = [\{b\}]$ . Since  $\mathbb{W} = [\{b\}]$  satisfies  $\neg\mathbf{K}a \wedge \mathbf{K}b$ , this is a G91-world view of  $\Gamma_2$ . If  $\mathbb{W} \models \mathbf{K}a \wedge \neg\mathbf{K}b$  instead, we get  $\Gamma_1^{\mathbb{W}} = \{\top \rightarrow a, \perp \rightarrow b\}$ , whose unique stable model is  $\{a\}$ . As a result,  $\mathbb{W} = [\{a\}]$  is also the other G91-world view of  $\Gamma_2$ . To see that there is not any other world views, note that  $\mathbb{W} \models \neg\mathbf{K}a \wedge \neg\mathbf{K}b$  implies that  $\Gamma_2^{\mathbb{W}} = \{\top \rightarrow a, \top \rightarrow b\}$  and  $\text{SM}[\Gamma_2^{\mathbb{W}}] = [\{a, b\}]$ . So this cannot a G91-world view. Similar, it can be checked that no world view can satisfy  $\mathbf{K}a \wedge \mathbf{K}b$ .  $\square$*

**Example 4.** *Take now the theory  $\Gamma_3 = \{\mathbf{K}a \rightarrow a\}$  corresponding to rule (3). If  $\mathbb{W} \models \mathbf{K}a$  we get  $\Gamma_3^{\mathbb{W}} = \{\top \rightarrow a\}$  and  $\text{SM}[\Gamma_3^{\mathbb{W}}] = \{a\}$  so  $\mathbb{W} = [\{a\}]$  is a G91-world view. If  $\mathbb{W} \not\models \mathbf{K}a$ , the reduct becomes  $\Gamma_3^{\mathbb{W}} = \{\perp \rightarrow a\}$ , a classical tautology with unique stable model  $\emptyset$ . As a result,  $\mathbb{W} = [\emptyset]$  is the second G91-world view of this theory.  $\square$*

### 2.3 Epistemic Specifications and Epistemic Splitting

In this section, we recall the formal definition of the Epistemic Splitting property. For the motivation of the interest of this property we refer to (Cabalar et al. 2018; Cabalar et al. 2019b). Let start by introducing a particular class of theories that correspond to the syntax of epistemic specifications or (*epistemic logic*) programs. Given a set of atoms  $S \subseteq \text{At}$ , by  $\neg S \stackrel{\text{def}}{=} \{\neg a \mid a \in S\}$  and  $\neg\neg S \stackrel{\text{def}}{=} \{\neg\neg a \mid a \in S\}$  we respectively denote the set resulting of prepend one or two occurrences of the default negation operator  $\neg$  to every atom in  $S$ . An *objective literal* is either an atom or a truth constant<sup>4</sup>, that is  $a \in \text{At} \cup \{\top, \perp\}$ , or the result of prepend one or two default negation,  $\neg a$ . By  $\text{Lit}^{obj} \stackrel{\text{def}}{=} \text{At} \cup \neg\text{At} \cup \neg\neg\text{At} \cup \{\top, \perp, \neg\top, \neg\perp, \neg\neg\top, \neg\neg\perp\}$  we denote the set of all objective literals. A *subjective literal* is any expression of the form  $\mathbf{K}l, \neg\mathbf{K}l$  or  $\neg\neg\mathbf{K}l$  with  $l \in \text{Lit}^{obj}$  any objective literal. A *literal* is either an objective or subjective literal.

A rule  $r$  is an implication of the form:

$$a_1 \vee \dots \vee a_n \leftarrow L_1 \wedge \dots \wedge L_m \tag{5}$$

with  $n \geq 0$  and  $m \geq 0$ , where each  $a_i \in \text{At}$  is an atom and each  $L_j$  a literal. The left hand disjunction of (5) is called the rule *head* and abbreviated as  $\text{Head}(r)$ . When

<sup>4</sup> For a simpler description of program transformations, we allow truth constants with their usual meaning.

$n = 0$ , it corresponds to  $\perp$  and  $r$  is called a *constraint*. The right hand side of (5) is called the rule *body* and abbreviated as  $Body(r)$ . As usual, we define  $Body^+(r)$  and  $Body^-(r)$  as the conjunction of all positive and negative literals in  $Body(r)$ , respectively, so that  $Body(r) \equiv Body^+(r) \wedge Body^-(r)$ . We further define  $Body_{obj}(r)$  and  $Body_{sub}(r)$  as the conjunction of all objective and subjective literals in  $Body(r)$ , respectively, so that  $Body(r) \equiv Body_{obj}(r) \wedge Body_{sub}(r)$ . We also define  $Body_y^x(r) \stackrel{\text{def}}{=} Body^x(r) \cap Body_y(r)$  with  $x \in \{+, -\}$  and  $y \in \{obj, sub\}$ . By abuse of notation, we will also use sometimes  $Body^x$ ,  $Body_y$  and  $Body_y^x$  as the set of literals occurring in those formulas. When  $m = 0$ , the body corresponds to  $\top$  and  $r$  is called a *fact* (in this case, the body and the arrow symbol are usually omitted). A rule is called *objective* if all literals occurring in it are objective. A *program*  $\Pi$  is a (possibly infinite) set of rules and an *objective program* is a program where all its rules are objective. We are now ready to recall the epistemic splitting property:

**Definition 3** (Epistemic splitting set). *A set of atoms  $U \subseteq At$  is said to be an epistemic splitting set of a program  $\Pi$  if for any rule  $r$  in  $\Pi$  one of the following conditions hold*

- (i)  $Atoms(r) \subseteq U$ ,
- (ii)  $(Atoms(Body_{obj}(r) \cup Head(r))) \cap U = \emptyset$

We define a splitting of  $\Pi$  as a pair  $\langle B_U(\Pi), T_U(\Pi) \rangle$  satisfying  $B_U(\Pi) \cap T_U(\Pi) = \emptyset$ ,  $B_U(\Pi) \cup T_U(\Pi) = \Pi$ , all rules in  $B_U(\Pi)$  satisfy (i) and all rules in  $T_U(\Pi)$  satisfy (ii).  $\square$

With respect to the original definition of splitting set, we can see that the condition for the top program  $Atoms(Head(r)) \cap U = \emptyset$  was replaced by the new condition (ii), which intuitively means that the top program may only refer to atoms  $U$  in the bottom through epistemic operators. Another observation is that the definition of  $B_U(\Pi)$  and  $T_U(\Pi)$  is kept non-deterministic in the sense that some rules can be arbitrarily included in one set or the other. These rules correspond to subjective constraints on atoms in  $U$ , since these are the only cases that may satisfy conditions (i) and (ii) simultaneously. Then, the idea is similar as in splitting a regular program: first we compute the world views of the bottom program  $B_U(\Pi)$  and for each one we compute the world views of the top program after simplifying in it the subjective literals fixed by the bottom part. Formally, given an epistemic splitting set  $U$  for a program  $\Pi$  and belief view  $\mathbb{W}$ , we define  $E_U(\Pi, \mathbb{W}) \stackrel{\text{def}}{=} T_U(\Pi)_{\mathbb{W}}^U$ , that is, we make the subjective reduct of the top with respect to  $\mathbb{W}$  and signature  $U$ .

**Definition 4.** *Given a semantics  $\mathcal{S}$ , a pair  $\langle \mathbb{W}_b, \mathbb{W}_t \rangle$  is said to be a  $\mathcal{S}$ -solution of  $\Pi$  with respect to an epistemic splitting set  $U$  if  $\mathbb{W}_b$  is a  $\mathcal{S}$ -world view of  $B_U(\Pi)$  and  $\mathbb{W}_t$  is a  $\mathcal{S}$ -world view of  $E_U(\Pi, \mathbb{W}_b)$ .  $\square$*

Notice that this definition depends on a particular semantics  $\mathcal{S}$  in the sense that each alternative semantics for epistemic specifications will define its own solutions for a given  $U$  and  $\Pi$ . In particular, in this paper, we will consider five instantiations of this Definition 4 with semantics  $\mathcal{S} \in \{G91, FAEEL, FEEL, EEL, AEEL\}$ . Besides the already mentioned G91, Founded Autoepistemic EEL (FAEEL) and Founded EEL (FEEL) semantics,<sup>5</sup> we will also consider the EEL and AEEL semantics from (Fariñas del Cerro et al. 2015).

<sup>5</sup> To avoid possible confusions between FAEEL, FEEL, AEEL and EEL, we will sometimes write them as Founded Autoepistemic EEL, Founded EEL, Autoepistemic EEL and EEL respectively, when they occur in the same sentence.

**Example 5** (Example 3 continued). *Back to our example, let now  $\Pi_4$  be the program consisting of rules (1)-(2) and (4). Then, we can see that  $U = \{a, b\}$  is an epistemic splitting set of  $\Pi_4$  and that it satisfies  $B_U(\Pi_4) = \{(1) - (2)\}$  and  $T_U(\Pi_4) = \{(4)\}$ . Furthermore, we have already seen that  $B_U(\Pi_4)$  corresponds to the theory  $\Gamma_2$  which has the following two G91-world views:  $[\{a\}]$  and  $[\{b\}]$ . Then, we can see that  $E_U(\Pi_4, [\{a\}]) = \{c \leftarrow \top\}$  has a unique G91-world view  $[\{c\}]$  and that  $E_U(\Pi_4, [\{b\}]) = \{c \leftarrow \perp\}$  has the unique G91-world view  $[\emptyset]$ . As a result, we have two G91-solutions of  $\Pi_4$  with respect to  $\{a, b\}$ :  $([\{a\}], [\{c\}])$  and  $([\{b\}], [\emptyset])$ . It is also easy to check that  $B_U(\Pi_4)$  has two G91-world views,  $[\{a, c\}]$  and  $[\{b\}]$  that can be obtained by composing the two above solutions.  $\square$*

In the general case, the world views for the global program are reconstructed by the following operation:

$$\mathbb{W}_b \sqcup \mathbb{W}_t = \{ I_b \cup I_t \mid I_b \in \mathbb{W}_b \text{ and } I_t \in \mathbb{W}_t \}$$

(remember that both the bottom and the top may produce multiple world views, depending on the program and the semantics we choose). For any set of atoms  $U \subseteq At$  and belief view  $\mathbb{W}$ , we also define the restriction of  $\mathbb{W}$  to  $U$  as  $\mathbb{W}_U \stackrel{\text{def}}{=} \{ I \cap U \mid I \in \mathbb{W} \}$ . Furthermore, we also define the complement of a set of atoms as  $\bar{U} \stackrel{\text{def}}{=} At \setminus U$ .

**Property 1** (Epistemic splitting). *A semantics  $\mathcal{S}$  satisfies epistemic splitting if for any epistemic splitting set  $U$  of any program  $\Pi$ :  $\mathbb{W}$  is a  $\mathcal{S}$ -world view of  $\Pi$  iff there is a  $\mathcal{S}$ -solution  $\langle \mathbb{W}_b, \mathbb{W}_t \rangle$  of  $\Pi$  with respect to  $U$  such that  $\mathbb{W} = \mathbb{W}_b \sqcup \mathbb{W}_t$ .  $\square$*

**Theorem 1.** *Semantics G91 satisfies epistemic splitting. Furthermore, if  $\mathbb{W}$  is a G91-world view of some program  $\Pi$  with respect to some splitting set  $U$ , then  $\langle \mathbb{W}_U, \mathbb{W}_{\bar{U}} \rangle$  is a G91-solution of  $\Pi$  and it satisfies that  $\mathbb{W} = \mathbb{W}_U \sqcup \mathbb{W}_{\bar{U}}$ .  $\square$*

Theorem 1 was proved in (Cabalar et al. 2019b, Main Theorem). Note that there, it is only stated that G91 satisfies epistemic splitting, however it is easy to see that the second part of the statement was proved as an auxiliary result inside the proof of that theorem. We decided to explicitly state this result as it will be useful for proving that FAEEL satisfies epistemic splitting.

### 2.4 Founded Autoepistemic Equilibrium Logic

We recall now the semantics of *Founded Autoepistemic Equilibrium Logic* (FAEEL) from (Cabalar et al. 2019a). The basic idea is an elaboration of the belief (or KD45) interpretation  $(\mathbb{W}, I)$  already seen but replacing belief sets by HT pairs. Thus, the idea of *belief view*  $\mathbb{W}$  is extended to a non-empty set of HT-interpretations  $\mathbb{W} = \{ \langle H_1, T_1 \rangle, \dots, \langle H_n, T_n \rangle \}$  and say that  $\mathbb{W}$  is *total* when  $H_i = T_i$  for all of them, coinciding with the form of belief views  $\mathbb{W} = \{ T_1, \dots, T_n \}$  we had so far. Similarly, a *belief interpretation* is now redefined as  $(\mathbb{W}, \langle H, T \rangle)$ , or simply  $(\mathbb{W}, H, T)$ , where  $\mathbb{W}$  is a belief view and  $\langle H, T \rangle$  stands for the real world, possibly not in  $\mathbb{W}$ . A belief interpretation  $(\mathbb{W}, H, T)$  is called *total* iff both  $\langle H, T \rangle$  and  $\mathbb{W}$  are total. Next, the satisfaction relation is defined as a combination of modal logic KD45 and HT. A belief interpretation  $\mathcal{I} = (\mathbb{W}, H, T)$  satisfies a formula  $\varphi$ , written  $\mathcal{I} \models \varphi$ , iff:

- $\mathcal{I} \not\models \perp$ ,
- $\mathcal{I} \models a$  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  $(\mathbb{W}^t, T) \models \psi_2$ , where<sup>6</sup>  $\mathbb{W}^t = \{T_i \mid \langle H_i, T_i \rangle \in \mathbb{W}\}$ .
- $\mathcal{I} \models \mathbf{K} \psi$  iff  $(\mathbb{W}, H_i, T_i) \models \psi$  for all  $\langle H_i, T_i \rangle \in \mathbb{W}$ .

A belief interpretation  $(\mathbb{W}, H, T)$  is called a *belief model* of a theory  $\Gamma$  iff  $(\mathbb{W}, H_i, T_i) \models \varphi$  for all HT-interpretation  $\langle H_i, T_i \rangle \in \mathbb{W} \cup \{\langle H, T \rangle\}$  and all  $\varphi \in \Gamma$ . Given theories  $\Gamma$  and  $\Gamma'$ , we write  $\Gamma \models \Gamma'$  when  $(\mathbb{W}, H, T) \models \Gamma$  implies  $(\mathbb{W}, H, T) \models \Gamma'$  for all belief interpretations. We write  $\Gamma \equiv \Gamma'$  iff  $\Gamma \models \Gamma'$  and  $\Gamma' \models \Gamma$ . Furthermore, when  $\Gamma$  or  $\Gamma'$  are singletons we may omit the brackets around their unique formula.

Recall that the negation of a formula  $\neg\varphi$  is defined as an abbreviation for the implication  $\varphi \rightarrow \perp$ . The following result is immediate from the above definition plus the persistence property proved in (Cabalar et al. 2019a) (Proposition 1) and explicitly states the evaluation of negation:

**Proposition 1.** *Given a belief interpretation  $\mathcal{I} = (\mathbb{W}, H, T)$  and a formula  $\varphi$ , it follows that  $\mathcal{I} \models \neg\varphi$  iff  $(\mathbb{W}^t, T) \not\models \varphi$ . □*

As recalled in Section 2.1, stable models correspond to a class of HT-models called equilibrium models, that is, total minimal models. Similarly, we define now equilibrium belief models as total minimal belief models with respect to the following order relation:

**Definition 5.** *We define the partial order  $\mathcal{I}' \preceq \mathcal{I}$  for belief interpretations  $\mathcal{I}' = (\mathbb{W}', H', T')$  and  $\mathcal{I} = (\mathbb{W}, H, T)$  when the following three conditions hold:*

- (i)  $T' = T$  and  $H' \subseteq H$ , and
- (ii) for every  $\langle H_i, T_i \rangle \in \mathbb{W}$ , there is some  $\langle H'_i, T_i \rangle \in \mathbb{W}'$ , with  $H'_i \subseteq H_i$ .
- (iii) for every  $\langle H'_i, T_i \rangle \in \mathbb{W}'$ , there is some  $\langle H_i, T_i \rangle \in \mathbb{W}$ , with  $H'_i \subseteq H_i$ . □

As usual,  $\mathcal{I}' \prec \mathcal{I}$  means  $\mathcal{I}' \preceq \mathcal{I}$  and  $\mathcal{I}' \neq \mathcal{I}$ .

**Definition 6.** *A total belief interpretation  $\mathcal{I} = (\mathbb{W}, T)$  is said to be an equilibrium belief model of some theory  $\Gamma$  iff  $\mathcal{I}$  is a belief model of  $\Gamma$  and there is no other belief model  $\mathcal{I}'$  of  $\Gamma$  such that  $\mathcal{I}' \prec \mathcal{I}$ . □*

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

**Definition 7.** *A total belief view  $\mathbb{W}$  is called an autoepistemic equilibrium model or FAEEL-world view of  $\Gamma$  iff:*

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

**Theorem 2** (Main Theorem in (Cabalar et al. 2019a)). *For any theory  $\Gamma$ , its FAEEL-world views are exactly its founded<sup>7</sup> G91-world views of  $\Gamma$ . □*

<sup>6</sup> Note that  $\mathbb{W}^t$  is a belief view as defined in Section 2.2.

<sup>7</sup> For space reasons we omit here the definition of *founded world view* and refer the reader to (Cabalar et al. 2019a). Intuitively, a world view is founded if all atoms that are true in all its belief set can be derived without cyclic references. If we omit there founded, we obtain that every FAEEL-world view is also a G91-world view, but not necessarily vice-versa.

### 3 Founded Epistemic Equilibrium Logic

Founded EEL is similar to Founded Autoepistemic EEL, but without the minimisation of knowledge. Technically, this makes Founded EEL 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 7). Note that, as mentioned in the introduction Founded EEL can be seen as the combination of the stable model semantics with the modal logic S5, while Founded Autoepistemic EEL would be the combination of the stable model semantics with Moore’s Autoepistemic Logic. In this sense, (i) is a direct consequence of the fact that Moore’s Autoepistemic Logic is defined in terms of modal logic KD45 instead of S5. In its turn, (ii) is a consequence of the fact that S5 is a monotonic logic, and thus Founded EEL do not need the autoepistemic fixpoint condition (Definition 7) that Founded Autoepistemic EEL inherits from Moore’s Autoepistemic Logic.

Formally, a belief view  $\mathbb{W}$  is called an *epistemic model* of a theory  $\Gamma$ , in symbols  $\mathbb{W} \models \Gamma$  iff  $(\mathbb{W}, H_i, T_i) \models \varphi$  for all HT-interpretation  $\langle H_i, T_i \rangle \in \mathbb{W}$  and all  $\varphi \in \Gamma$ . Given theories  $\Gamma$  and  $\Gamma'$ , we write  $\Gamma \models_{\text{FEEL}} \Gamma'$  when  $(\mathbb{W}, H, T) \models \Gamma$  implies  $(\mathbb{W}, H, T) \models \Gamma'$  for all belief interpretations. We write  $\Gamma \equiv_{\text{FEEL}} \Gamma'$  iff  $\Gamma \models_{\text{FEEL}} \Gamma'$  and  $\Gamma' \models_{\text{FEEL}} \Gamma$ . As above, when  $\Gamma$  or  $\Gamma'$  are singletons we may omit the brackets around their unique formula.

**Proposition 2** (Persistence).  $\mathbb{W} \models \varphi$  implies  $\mathbb{W}^t \models \varphi$ . □

*Proof.* Follows directly from Proposition 1 in (Cabalar et al. 2019a). □

The following order relation adapts Definition 5 to the case of belief views.

**Definition 8.** Given belief views  $\mathbb{W}_1$  and  $\mathbb{W}_2$ , we write  $\mathbb{W}_1 \preceq \mathbb{W}_2$  iff the following two condition hold:

- (i) for every  $\langle H_2, T \rangle \in \mathbb{W}_2$ , there is some  $\langle H_1, T \rangle \in \mathbb{W}_1$ , with  $H_1 \subseteq H_2$ .
- (ii) for every  $\langle H_1, T \rangle \in \mathbb{W}_1$ , there is some  $\langle H_2, T \rangle \in \mathbb{W}_2$ , with  $H_1 \subseteq H_2$ .

As usual, we write  $\mathbb{W}_1 \prec \mathbb{W}_2$  iff  $\mathbb{W}_1 \preceq \mathbb{W}_2$  and  $\mathbb{W}_1 \neq \mathbb{W}_2$ . □

Then, equilibrium epistemic models are defined as usual:

**Definition 9.** A total epistemic model  $\mathbb{W}$  of a theory  $\Gamma$  is said to be an epistemic equilibrium model or FEEL-world view iff there is no other epistemic model  $\mathbb{W}'$  of  $\Gamma$  such that  $\mathbb{W}' \prec \mathbb{W}$ . □

The following observation establishes a relation between (equilibrium) belief models and (equilibrium) epistemic models similar to the existent between the standard modal logics KD45 and S5. Recall that belief models correspond to the logical product of HT and KD45 while epistemic models come from the logical product of HT and S5.

**Observation 1.** For any theory  $\Gamma$  and belief interpretation  $\mathcal{I} = \langle \mathbb{W}, H, T \rangle$ , the following statements hold:

- (i) If  $\mathcal{I}$  is a belief model of  $\Gamma$ , then  $\mathbb{W}$  is a epistemic model of  $\Gamma$ , and
- (ii) If  $\mathcal{I}$  is an equilibrium belief model of  $\Gamma$ , then  $\mathbb{W}$  is a equilibrium epistemic model of  $\Gamma$ . □

From this observation the following relation between autoepistemic equilibrium models and epistemic equilibrium models can be established:

**Theorem 3.** *Every FAEEL-world view of any theory  $\Gamma$  is also an FEEL-world view of  $\Gamma$ .*  $\square$

In general the converse does not necessary holds as illustrated by the following example:

**Example 6.** *Consider the program  $\Pi_1$  consisting of the single rule  $a \vee b$ . This program has two stable models,  $\{a\}$  and  $\{b\}$  and accordingly a unique FAEEL-world view  $[\{a\}, \{b\}]$  which agrees with its unique G91-world view. On the other hand, this program has two extra FEEL-world views that are not FAEEL-world views:  $[\{a\}]$  and  $[\{b\}]$ . These two extra models are clearly not well-justified in an epistemic sense as they respectively satisfy  $\mathbf{K}a$  and  $\mathbf{K}b$  with no evidence for that conclusion. Finally, to see the difference between FEEL and modal logic S5, note that  $[\{a, b\}]$  is also an S5 model which is neither a FEEL nor a FAEEL-world view.*  $\square$

This example also illustrates that the difference between FAEEL and FEEL can be formalised in terms of the *supra-ASP* property introduced in (Cabalar et al. 2019b) and recalled below: FAEEL satisfies supra-ASP (Cabalar et al. 2019a, Proposition 3) while FEEL does not.

**Property 2** (Supra-ASP). *A semantics  $\mathcal{S}$  satisfies supra-ASP if for any objective program  $\Pi$  either  $\Pi$  has a unique  $\mathcal{S}$ -world view  $\mathbb{W} = \text{SM}[\Pi] \neq \emptyset$  or both  $\text{SM}[\Pi] = \emptyset$  and  $\Pi$  has no  $\mathcal{S}$ -world view at all.*  $\square$

On the other hand, the following result shows that we can obtain FAEEL-world views as the intersection of FEEL and G91-world views:

**Theorem 4.** *For any theory  $\Gamma$ , a belief view  $\mathbb{W}$  is a FAEEL-world view iff (i)  $\mathbb{W}$  is a FEEL-world view and (ii)  $\mathbb{W}$  is a G91-world view.*  $\square$

**Example 7** (Example 4 continued). *Back to  $\Gamma_3 = \{\mathbf{K}a \rightarrow a\}$ , recall that this theory has two G91-world views:  $[\emptyset]$  and  $[\{a\}]$ . It is easy to see that  $[\emptyset]$  is a FEEL-world view and, from Theorem 4, this implies that this is also a FAEEL-world view. Note that there is no smaller belief view than  $[\emptyset]$ , so being a model is enough to show that is a FEEL-world view. Note that, by using Theorem 4, is much easier to see that  $[\emptyset]$  is a FAEEL-world view than directly using its definition since we would also need to check no other  $T \notin [\emptyset]$  satisfies  $([\emptyset], T) \in \text{EQB}[\Gamma_3]$ . In this case, the only possibility is  $([\emptyset], \{a\})$  which fails because there is a smaller belief model  $([\emptyset], \emptyset, \{a\})$  satisfying  $\mathbf{K}a \rightarrow a$ . On the other hand, we can see that  $[\{a\}]$  is not a FEEL-world view and, thus neither FAEEL-world view, because  $[\emptyset, \{a\}]$  also satisfies  $\Gamma_3$ . In this case, it is also easier to use Theorem 4 than using the FAEEL definition:  $[\{a\}]$  is not a FAEEL-world view because  $\mathcal{I}' = (\{[\{a\}], \{a\}\}) \notin \text{EQB}[\Gamma_3]$  and this is the case because the smaller interpretation  $\mathcal{I}'' = ([\emptyset, \{a\}], \{a\}, \{a\})$  also satisfies  $\Gamma_3$ . In particular, note that  $\mathcal{I}'' \not\models \mathbf{K}a$  and, thus, clearly satisfies  $\mathbf{K}a \rightarrow a$ .*  $\square$

Example 7 illustrates how Theorem 4 can be used to find FAEEL-world views using FEEL. In general, this is much easier than directly applying their definition because belief views are simpler than belief interpretations and because the autoepistemic fixpoint can be checked independently using the G91-semantics.

### 4 Epistemic Splitting in Founded (Auto)Epistemic Equilibrium Logic

Let us now study the epistemic splitting property in FEEL and FAEEL. Let us start by stating a result analogous to Theorem 1, but this time for FEEL.

**Theorem 5** (Epistemic splitting in FEEL). *FEEL satisfies the epistemic splitting property. Furthermore, if  $\mathbb{W}$  is a FEEL-world view of some program  $\Pi$  with respect to some splitting set  $U$  then  $\langle \mathbb{W}_U, \mathbb{W}_{\bar{U}} \rangle$  is a FEEL-solution of  $\Pi$  and it satisfies that  $\mathbb{W} = \mathbb{W}_U \sqcup \mathbb{W}_{\bar{U}}$ .*  $\square$

The proof of Theorem 5 is based in the following auxiliary results whose proof can be found in the supplementary material.

**Proposition 3.** *Let  $U \subseteq \text{At}$  be some set of atoms and  $\Pi = \Pi_1 \cup \Pi_2$  be a program such that  $\text{Atoms}(\Pi_1) \subseteq U$  and  $\text{Atoms}(\Pi_2) \subseteq \bar{U}$ . Then, any belief view  $\mathbb{W}$  is an FEEL-world view of  $\Pi$  iff  $\mathbb{W}_U$  is an FEEL-world view of  $\Pi_1$  and  $\mathbb{W}_{\bar{U}}$  is an FEEL-world view of  $\Pi_2$ .*  $\square$

Intuitively, Proposition 3 says that, if we can split a program in a way that its two halves do not share atoms in common, then we can compute the world views of the whole program by combining the world views of each half. Furthermore, the following result shows that we can check whether a belief view is an FEEL-world view by checking instead that that belief view is a FEEL-world view of a program obtained by simplifying the subjective literals in the rules of the top part accordingly to the belief view.

**Proposition 4.** *Let  $\Pi$  be a program with epistemic splitting set  $U \subseteq \text{At}$ . Then, any belief view  $\mathbb{W}$  is an FEEL-world view of  $\Pi$  iff  $\mathbb{W}$  is an FEEL-world view of  $B_U(\Pi) \cup E_U(\Pi, \mathbb{W})$ .*  $\square$

We can now join Propositions 3 and 4 to show the following rewriting of epistemic splitting.

**Proposition 5.** *Given any program  $\Pi$  and epistemic splitting set  $U$  of  $\Pi$ , a belief view  $\mathbb{W}$  is an FEEL-world view of  $\Pi$  iff  $\mathbb{W}_U$  is an FEEL-world view of  $B_U(\Pi)$  and  $\mathbb{W}_{\bar{U}}$  is an FEEL-world view of  $\mathbb{W}_{\bar{U}}$  of  $E_U(\Pi, \mathbb{W}_U)$ .*  $\square$

*Proof.* First note that, from Proposition 4, it follows that  $\mathbb{W}$  is an FEEL-world view of  $\Pi$  iff  $\mathbb{W}$  is an FEEL-world view of  $B_U(\Pi) \cup E_U(\Pi, \mathbb{W})$ . Furthermore, we have  $E_U(\Pi, \mathbb{W}) = E_U(\Pi, \mathbb{W}_U)$ . Therefore, we immediately can see that  $\mathbb{W}$  is an FEEL-world view of  $\Pi$  iff  $\mathbb{W}$  is an FEEL-world view of  $B_U(\Pi) \cup E_U(\Pi, \mathbb{W}_U)$ . Furthermore, by construction we have  $\text{Atoms}(B_U(\Pi)) \subseteq U$  and  $\text{Atoms}(E_U(\Pi, \mathbb{W}_U)) \subseteq \bar{U}$  and, from Proposition 3, this implies that the latter holds iff  $\mathbb{W}_U$  is an FEEL-world view of  $B_U(\Pi)$  and  $\mathbb{W}_{\bar{U}}$  is an FEEL-world view of  $E_U(\Pi, \mathbb{W}_U)$ .  $\square$

**Proof of Theorem 5.** Assume first that  $\mathbb{W}$  is an equilibrium epistemic model of  $\Pi$ . Then, from Proposition 5, it follows that  $\mathbb{W}_U$  is an FEEL-world view of  $B_U(\Pi)$  and  $\mathbb{W}_{\bar{U}}$  is an FEEL-world view of  $\mathbb{W}_{\bar{U}}$  of  $E_U(\Pi, \mathbb{W}_U)$  and it is easy to check that  $\mathbb{W} = \mathbb{W}_U \sqcup \mathbb{W}_{\bar{U}}$ . The other way around, assume there are FEEL-world views  $\mathbb{W}_b$  of  $B_U(\Pi)$  and  $\mathbb{W}_t$  of  $E_U(\Pi, \mathbb{W}_b)$  and let  $\mathbb{W} = \mathbb{W}_b \sqcup \mathbb{W}_t$ . Note that  $\text{Atoms}(B_U(\Pi)) \subseteq U$  implies that every interpretation  $T \in \mathbb{W}_b$  satisfies  $T \subseteq U$  and that  $\text{Atoms}(E_U(\Pi, \mathbb{W}_b)) \subseteq \bar{U}$  implies that

every  $T \in \mathbb{W}_t$  satisfies  $T \subseteq \overline{U}$ . Hence, it follows that  $\mathbb{W}_b = \mathbb{W}_U$  and  $\mathbb{W}_t = \mathbb{W}_{\overline{U}}$  and the result follows directly from Proposition 5.  $\square$

We can now use Theorem 5 in combination with Theorems 1 and 4 to show that FAEEL also satisfies epistemic splitting.

**Main Theorem** (Epistemic splitting in FAEEL). *FAEEL satisfies the epistemic splitting property. Furthermore, if  $\mathbb{W}$  is a FAEEL-world view of some program  $\Pi$  with respect to some splitting set  $U$  then  $\langle \mathbb{W}_U, \mathbb{W}_{\overline{U}} \rangle$  is a FAEEL-solution of  $\Pi$  and it satisfies that  $\mathbb{W} = \mathbb{W}_U \sqcup \mathbb{W}_{\overline{U}}$ .*  $\square$

*Proof.* Assume first that there is FAEEL-solution  $\langle \mathbb{W}_b, \mathbb{W}_t \rangle$  of some program  $\Pi$  with respect to some splitting set  $U$  and let  $\mathbb{W} = \mathbb{W}_b \sqcup \mathbb{W}_t$ . By definition, this implies that  $\mathbb{W}_b$  is a FAEEL-world view of  $B_U(\Pi)$  which, from Theorem 4, implies that  $\mathbb{W}_b$  is both a FEEL and a G91-world view of  $B_U(\Pi)$ . Similarly, we can see that  $\mathbb{W}_t$  is both a FEEL and a G91-world view of  $E_U(\Pi, \mathbb{W}_b)$  and, thus, that  $\langle \mathbb{W}_b, \mathbb{W}_t \rangle$  is both a FEEL and a G91-solution of  $\Pi$  with respect  $U$ . From Theorems 1 and 5, these two facts respectively imply that  $\mathbb{W}$  is both a FEEL and a G91-world view of  $\Pi$  which, from Theorem 4 again, implies that  $\mathbb{W}$  is a FAEEL-world view of  $\Pi$ .

The other way around is analogous. Assume now that  $\mathbb{W}$  is a FAEEL-world view of some program  $\Pi$  with splitting set  $U$ . Then, from Theorem 4, it follows that  $\mathbb{W}$  is both a FEEL and a G91-world view. From Theorems 1 and 5, these two facts respectively imply that  $\langle \mathbb{W}_U, \mathbb{W}_{\overline{U}} \rangle$  is a FEEL and a G91-solution of  $\Pi$  with respect to  $U$  and  $\mathbb{W} = \mathbb{W}_U \sqcup \mathbb{W}_{\overline{U}}$ . Finally, from Theorem 4 again, this implies that  $\langle \mathbb{W}_U, \mathbb{W}_{\overline{U}} \rangle$  is also a FAEEL-solution.  $\square$

It is interesting to note that for any semantics that satisfies epistemic splitting, thus FAEEL and G91, constraints indented to remove world views are well-behaved:

**Property 3** (Subjective constraint monotonicity). *A semantics satisfies subjective constraint monotonicity if, for any epistemic program  $\Pi$  and any subjective constraint  $r$ ,  $W$  is a world view of  $\Pi \cup \{r\}$  iff both  $W$  is a world view of  $\Pi$  and  $W \models r$ .*  $\square$

**Theorem 6** (Theorem 2 in Cabalar et al. 2019b). *Epistemic splitting implies subjective constraint monotonicity.*  $\square$

Furthermore, this property also guarantees that, for *epistemically stratified* programs, these semantics have at most a unique world view (see Theorem 1 in Cabalar et al. 2019b). Another interesting property, not included in (Cabalar et al. 2019b), is that any semantics that satisfies epistemic splitting and the *supra-ASP* property, necessary coincides with the G91-semantics for the class of epistemic stratified programs. Before tackling the notion of epistemic stratification, let us recall the epistemic dependence relation among atoms in a program  $\Pi$  so that  $dep(a, b)$  is true iff there is a rule  $r \in \Pi$  such that  $a \in Atoms(Head(r) \cup Body_{obj}(r))$  and  $b \in Atoms(Body_{sub}(r))$ .

**Definition 10.** *We say that an epistemic program  $\Pi$  is epistemically stratified if we can assign an integer mapping  $\lambda : At \rightarrow \mathbb{N}$  to each atom such that  $\lambda(a) > \lambda(b)$  for any pair of atoms  $a, b$  satisfying  $dep(a, b)$ .*  $\square$

**Theorem 7.** *Given any two semantics  $\mathcal{S}$  and  $\mathcal{S}'$  that satisfy the epistemic splitting and supra-ASP properties and an epistemically stratified program  $\Pi$ , one of following two condition hold:*

- (i)  $\Pi$  has neither  $\mathcal{S}$ -world view nor  $\mathcal{S}'$ -world view, or
- (ii)  $\Pi$  has exactly one  $\mathcal{S}$  and one  $\mathcal{S}'$ -world view  $\mathbb{W}$  which is the same in both semantics.

*Proof.* The proof follows by induction in the number of layers induced by the stratifications. Note that, if a program has a unique layer, then it must be objective and, thus, the result follows directly from the supra-ASP property. Otherwise, let  $a \in At$  such that there is no  $b \in At$  with  $\lambda(a) < \lambda(b)$  and let  $U = \{ c \in At \mid \lambda(c) < \lambda(a) \}$ . Then,  $U$  is an splitting set of  $\Pi$  and the epistemic splitting property tell us a belief view  $\mathbb{W}$  is a  $\mathcal{S}$ -world view of  $\Pi$  iff there are  $\mathcal{S}$ -world views  $\mathbb{W}_b$  of  $B_U(\Pi)$  and  $\mathbb{W}_t$  of  $E_U(\Pi, \mathbb{W}_b)$ . Furthermore,  $B_U(\Pi)$  and  $E_U(\Pi, \mathbb{W}_b)$  have less layers than  $\Pi$  so, by induction hypothesis, this holds iff there are  $\mathcal{S}'$ -world views  $\mathbb{W}_b$  of  $B_U(\Pi)$  and  $\mathbb{W}_t$  of  $E_U(\Pi, \mathbb{W}_b)$  iff  $\mathbb{W}$  is a  $\mathcal{S}'$ -world view of  $\Pi$ . □

**Corollary 1.** *For epistemically stratified programs, FAEEL and G91-world views coincide.* □

Note that Corollary 1 does not apply to FEEL because this semantics does not satisfy supra-ASP as illustrated by Example 6. Recall also that, from Proposition 2 in (Cabalar et al. 2019a), we already knew that, for programs where all occurrences of  $\mathbf{K}$  are in the scope of negation, FAEEL and G91-world views coincide. Corollary 1 enlarges the class programs in which FAEEL and G91 coincide by including all those that are epistemically stratified. This immediately arises the question whether FAEEL and G91 also coincide for programs without positive cycles involving epistemic literals. The following result shows that this is indeed the case. Formally, we define the positive epistemic dependence relation among atoms in a program  $\Pi$  so that  $dep^+(a, b)$  is true iff there is a rule  $r \in \Pi$  such that  $a \in Atoms(Head(r) \cup Body_{obj}(r))$  and  $b \in Atoms(Body_{sub}^+(r))$ .

**Definition 11.** *We say that an epistemic program  $\Pi$  is epistemically tight if we can assign an integer mapping  $\lambda : At \rightarrow \mathbb{N}$  to each atom such that  $\lambda(a) > \lambda(b)$  for any pair of atoms  $a, b$  satisfying  $dep^+(a, b)$ .* □

**Definition 12.** *Given an epistemic theory  $\Gamma$  and a belief view  $\mathbb{W}$ , its negatively subjective reduct, in symbols  $\Gamma^{\mathbb{W}}$ , is obtained by replacing each maximal subjective formula of the form  $\neg \mathbf{K} \varphi$  by  $\top$  if  $\mathbb{W} \not\models \mathbf{K} \varphi$ ; or by  $\perp$  otherwise.* □

**Proposition 6.** *Given a theory  $\Gamma$  and, a total belief view  $\mathbb{W}$  is FAEEL-world view of  $\Gamma$  iff  $\mathbb{W}$  is a FAEEL-world view of  $\Gamma^{\mathbb{W}}$ .*

**Theorem 8.** *For epistemically tight programs, FAEEL and G91-world views coincide.* □

*Proof.* For any belief view  $\mathbb{W}$  and epistemically tight program  $\Pi$  it follows that  $\mathbb{W}$  is FAEEL-world view of  $\Pi$  iff  $\mathbb{W}$  is a FAEEL-world view of  $\Pi^{\mathbb{W}}$  (Proposition 6). Furthermore, it is easy to see that, since  $\Pi$  is epistemically tight,  $\Pi^{\mathbb{W}}$  is also epistemically tight. Moreover, for every  $r \in \Pi^{\mathbb{W}}$  and every  $L \in Body_{sub}(r)$ , we can check that  $L \in Body^+(r)$ . That is, there are no subjective literals in the scope of negation and, thus,  $\Pi^{\mathbb{W}}$  being

epistemically tight implies that  $\Pi^{\mathbb{W}}$  is epistemically stratified. Then, the result follows directly from Corollary 1.  $\square$

It is also worth to mention that, as observed in (Cabalar et al. 2019a), it is possible to obtain Moore's Autoepistemic Logic from FAEEL simply by adding the exclude middle axiom  $p \vee \neg p$  for every atom in the signature. Note that, if the original program was stratified, augmenting it with these formulas does not change this property. As a result we obtain the following corollary:

**Corollary 2.** *Any epistemically stratified program  $\Pi$  has at most one Moore's autoepistemic extension.*  $\square$

A similar result was originally proved in (Gelfond 1987, Theorem 4). Note that the class of programs considered stratified by us is slightly broader than the one used in (Gelfond 1987): constraints are allowed in every strata and no distinction is made between positive and negative objective literals. The price to pay is that Corollary 2 does not ensure the existence of an extension.

## 5 Related work

As mentioned in the introduction, the search for a “satisfactory” semantics for epistemic logic programs has leave us with a variety of semantics (Gelfond 1991; Wang and Zhang 2005; Truszczyński 2011; Gelfond 2011; Fariñas del Cerro et al. 2015; Kahl et al. 2015; Shen and Eiter 2017; Cabalar et al. 2019a). Among this Epistemic Equilibrium Logic (EEL; Fariñas del Cerro et al. 2015) is very similar to Founded EEL in the sense that it is also defined as a combination of Equilibrium Logic and the modal logic S5. There are some slight differences though, and as the name suggest Founded EEL satisfies the founded property defined (Cabalar et al. 2019a) while EEL does not. In fact, EEL can be characterised by a particular class of belief views that we call here *simple*:

**Definition 13.** *We say that a belief view  $\mathbb{W}$  is simple iff for any  $\langle H, T \rangle \in \mathbb{W}$  and  $\langle H', T \rangle \in \mathbb{W}$ , we have  $H = H'$ . A total belief view  $\mathbb{W}$  is called an EEL-world view of a theory  $\Gamma$  iff  $\mathbb{W}$  is a belief model of  $\Gamma$  and there is no simple belief model  $\mathbb{W}'$  of  $\Gamma$  satisfying  $\mathbb{W} = (\mathbb{W}')^t$  and  $H \subset T$  for some  $\langle H, T \rangle \in \mathbb{W}'$ .*  $\square$

It is easy to see that Definition 13 is just a rephrasing of epistemic equilibrium models as defined in (Fariñas del Cerro et al. 2015) by using the notation of this paper.

**Proposition 7.** *A total belief view  $\mathbb{W}$  is an EEL-model of a theory  $\Gamma$  iff  $\mathbb{W}$  is a belief model of  $\Gamma$  and there is no simple belief model  $\mathbb{W}'$  of  $\Gamma$  such that  $\mathbb{W}' \prec \mathbb{W}$ .*

**Theorem 9.** *Every FAEEL and FEEL-world view of any theory  $\Gamma$  is also an EEL-world view.*  $\square$

*Proof.* Note that that every total belief view is simple, though there are non-total belief views that are not simple. Then, the result follows directly from Proposition 7 and Theorem 3.  $\square$

As illustrated by the following example, in general, the converse of Theorem 9 does not hold.

**Example 8.** Take the epistemic logic program:

$$a \vee b \qquad a \leftarrow \mathbf{K} b \qquad b \leftarrow \mathbf{K} a \qquad (\Pi_2)$$

whose unique epistemic equilibrium model is  $\mathbb{W} = [\{a\}, \{b\}]$ . Note that  $\mathbb{W}' = [\{a, b\}]$  is not an epistemic equilibrium model because  $\mathbb{W}'' \prec \mathbb{W}'$  with  $\mathbb{W}'' = [\langle \{a\}, \{a, b\} \rangle, \langle \{b\}, \{a, b\} \rangle]$ . However,  $\mathbb{W}'$  is an EEL-model. Note that  $\mathbb{W}''$  is not a simple belief view and, thus, cannot be used as a witness to show that  $\mathbb{W}'$  is not an EEL-model. On the other hand, the simple belief views  $[\langle \{a\}, \{a, b\} \rangle]$ ,  $[\langle \{b\}, \{a, b\} \rangle]$ , and  $[\langle \emptyset, \{a, b\} \rangle]$  are not models of this program.  $\square$

Interestingly, it can be shown that EEL also satisfies epistemic splitting.

**Theorem 10** (Epistemic splitting in EEL). *EEL satisfies the epistemic splitting property. Furthermore, if  $\mathbb{W}$  is a EEL-world view of some program  $\Pi$  with respect to some splitting set  $U$  then  $(\mathbb{W}_U, \mathbb{W}_{\bar{U}})$  is a EEL-solution of  $\Pi$  and it satisfies that  $\mathbb{W} = \mathbb{W}_U \sqcup \mathbb{W}_{\bar{U}}$ .*  $\square$

The proof of Theorem 10 is analogous to the proof of Theorem 5 just taking into account that now we have to restrict ourselves to simple interpretations. Note that, in general, EEL does not satisfy supra-ASP. Example 6 can be used to illustrate this statement and, in fact, the program in this example has exactly the same EEL-world views as FEEL-world views. For this reason (Fariñas del Cerro et al. 2015) also included a selection of EEL-world views called AEEL-world views in a similar spirit as how FAEEL-world views are a selection of FEEL-world views. However, it has been shown in (Cabalar et al. 2018; Cabalar et al. 2019b) that AEEL does not satisfy epistemic splitting. Theorem 10 sheds more light into this issue by showing that it is not the EEL logic, but the selection of AEEL-world views, what breaks the splitting property. In this sense, it would be possible to define AEEL-world views in an alternative way as the intersection of EEL-world views and G91-world views and obtain yet another semantics that satisfy epistemic splitting. Note though that this alternative semantics (EEL+G91) would not satisfy the foundedness property introduced in (Cabalar et al. 2019a). In fact, a variation of Example 8, obtained by adding the constraint  $\perp \leftarrow \neg \mathbf{K} a$  to program  $\Pi_2$ , was used (Cabalar et al. 2019a) to show that, among others, AEEL does not satisfy the foundedness property. This same example can also be used to show that EEL and EEL+G91 do not satisfy it.

To summarise the state of the art, let us recall the remaining property introduced in (Cabalar et al. 2018; Cabalar et al. 2019b):

**Property 4** (Supra-S5). *A semantics satisfies supra-S5 when for every world view  $W$  of an epistemic program  $\Pi$  and for every  $I \in W$ ,  $(W, I) \models \Pi$ .*  $\square$

Table 1 summarises the known results for different semantics with respect to Properties 1-4 plus *foundedness*. Recall that, intuitively, *foundedness* means that a semantics is free of self-supported world view. For space reasons, we refer to (Cabalar et al. 2019a) for a formal definition. Recall also that (Wang and Zhang 2005; Truszczyński 2011) extended the semantics of G91 to arbitrary theories in different ways, but for the class of logic programs both of them agree with G91. For this reason, we will refer to both of them just as G91 in the table. Counterexamples for the non satisfied properties can be found in (Kahl and Leclerc 2018; Cabalar et al. 2018; Cabalar et al. 2019b; Cabalar et al. 2019a) and in this paper for the case of EEL and FEEL not satisfying Supra-ASP. Proofs

	G91	G11	EEL	AEL	EEL+G91	K15	S17	FEEL	FAEEL
Supra-S5	✓	✓	✓	✓	✓	✓	✓	✓	✓
Supra-ASP	✓	✓		✓	✓	✓	✓		✓
Subjective constraint monotonicity	✓	✓	✓		✓			✓	✓
Splitting	✓		✓		✓			✓	✓
Foundness								✓	✓

Table 1. *Summary of properties in different semantics.*

for the satisfied properties can be found in (Cabalar et al. 2018; Cabalar et al. 2019b; Cabalar et al. 2019a) and in this paper. To complete the table, the following result shows that, despite not satisfying epistemic splitting, the semantics proposed by Gelfond (2011) does satisfy subjective constraining monotonicity.

**Proposition 8.** *The semantics defined by Gelfond (2011) satisfies subjective constraint monotonicity.* □

## 6 Conclusions

We have shown that Founded Autoepistemic EEL satisfies the *epistemic splitting*, a desirable property for epistemic logic programs that, among previous semantics, was known to be satisfied only by G91. On the other hand, it is well-known that the G91 semantics suffers from self-supported world views, something that was proved to be not the case for Founded Autoepistemic EEL in (Cabalar et al. 2019a). In this sense, Founded Autoepistemic EEL is the first semantics whose world views are not self-supported and that satisfies epistemic splitting. Furthermore, we have shown that, for epistemic tight programs (those not containing cycles involving positive epistemic literals), both G91 and Founded Autoepistemic EEL coincide. This means that Founded Autoepistemic EEL corrects the problem with self-supported world views present in G91 without introducing further variations that are unrelated to this problem.

In addition, we have introduced Founded EEL, a logic which can be considered as a combination of the stable models semantics and the modal logic S5, and an alternative characterisation of Founded Autoepistemic EEL-world views in terms of Founded EEL and G91. This alternative characterisation may help us to further study properties of Founded Autoepistemic EEL and, in fact, it already has been used to prove the epistemic splitting property, strengthen the relation between Founded Autoepistemic EEL and G91, and also to study the relation with the Epistemic Equilibrium Logic introduced by Fariñas del Cerro et al. (2015).

## Supplementary material

To view supplementary material for this article, please visit <https://doi.org/10.1017/S1471068419000127>.

## References

- CABALAR, P., FANDINNO, J., AND FARIÑAS DEL CERRO, L. 2018. Splitting Epistemic Logic Programs. *arXiv e-prints*, arXiv:1812.08763.
- CABALAR, P., FANDINNO, J., AND FARIÑAS DEL CERRO, L. 2019a. Founded world views with autoepistemic equilibrium logic. In *LPNMR*. Lecture Notes in Computer Science (to appear). Springer.
- CABALAR, P., FANDINNO, J., AND FARIÑAS DEL CERRO, L. 2019b. Splitting epistemic logic programs. In *LPNMR*. Lecture Notes in Computer Science (to appear). Springer.
- FARIÑAS DEL CERRO, L., HERZIG, A., AND SU, E. I. 2015. Epistemic equilibrium logic. In *Proc. of the Intl. Joint Conference on Artificial Intelligence (IJCAI'15)*. AAAI Press, 2964–2970.
- GELDER, A. V., ROSS, K. A., AND SCHLIPF, J. S. 1991. The well-founded semantics for general logic programs. *J. ACM* 38, 3, 620–650.
- GELFOND, M. 1987. On stratified autoepistemic theories. In *AAAI*. Morgan Kaufmann, 207–211.
- GELFOND, M. 1991. Strong introspection. In *Proceedings of the AAAI Conference*, T. L. Dean and K. McKeown, Eds. Vol. 1. AAAI Press/The MIT Press, 386–391.
- GELFOND, M. 2011. New semantics for epistemic specifications. In *LPNMR*. Lecture Notes in Computer Science, vol. 6645. Springer, 260–265.
- GELFOND, M. AND LIFSCHITZ, V. 1988. The stable model semantics for logic programming. In *Proc. of the 5th Intl. Conference on Logic Programming (ICLP'88)*. 1070–1080.
- GELFOND, M. AND PRZYMUSINSKA, H. 1992. On consistency and completeness of autoepistemic theories. *Fundam. Inform.* 16, 1, 59–92.
- KAHL, P., WATSON, R., BALAI, E., GELFOND, M., AND ZHANG, Y. 2015. The language of epistemic specifications (refined) including a prototype solver. *Journal of Logic and Computation*.
- KAHL, P. T. AND LECLERC, A. P. 2018. Epistemic logic programs with world view constraints. In *ICLP (Technical Communications)*. OASICS, vol. 64. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 1:1–1:17.
- LECLERC, A. P. AND KAHL, P. T. 2018. Epistemic logic programs with world view constraints. In *Technical communication, 34th International Conference on Logic Programming (ICLP'2018)*.
- LEONE, N., RULLO, P., AND SCARCELLO, F. 1997. Disjunctive stable models: Unfounded sets, fixpoint semantics, and computation. *Inf. Comput.* 135, 2, 69–112.
- LIFSCHITZ, V. AND TURNER, H. 1994. Splitting a logic program. In *Proc. of the Intl. Conference on Logic Programming (ICLP'94)*. MIT Press, 23–37.
- MOORE, R. C. 1985. Semantical considerations on nonmonotonic logic. *Artif. Intell.* 25, 1, 75–94.
- SHEN, Y. AND EITER, T. 2017. Evaluating epistemic negation in answer set programming (extended abstract). In *Proc. of the Intl. Joint Conference on Artificial Intelligence (IJCAI'17)*. 5060–5064.
- TRUSZCZYŃSKI, M. 2011. Revisiting epistemic specifications. In *Logic Programming, Knowledge Representation, and Nonmonotonic Reasoning*. Lecture Notes in Computer Science, vol. 6565. Springer, 315–333.
- WANG, K. AND ZHANG, Y. 2005. Nested epistemic logic programs. In *LPNMR*. Lecture Notes in Computer Science, vol. 3662. Springer, 279–290.
- WATSON, R. 2000. A splitting set theorem for epistemic specifications. *CoRR: Proceedings of the 8th International Workshop on Non-Monotonic Reasoning, NMR 2000 cs.AI/0003038*.