

# Neighborhood Semantics for Logics of Unknown Truths and False Beliefs

David R. Gilbert and Giorgio Venturi\*

## Abstract

This article outlines a semantic approach to the logics of unknown truths, and the logic of false beliefs, using neighborhood structures, giving results on soundness, completeness, and expressivity. Relational semantics for the logics of unknown truths are also addressed, specifically the conditions under which sound axiomatizations of these logics might be obtained from their normal counterparts, and the relationship between reflexive insensitive logics (RI-logics) and logics containing the provability operator  $\Box$  as the primary modal operator.

## Introduction

This paper investigates the use of neighborhood frames as a semantics for the logics of unknown truths and related non-normal epistemic systems, including the logic of false beliefs. However, the interest in these logics is not restricted to epistemic settings. As such, this paper is perhaps better viewed as a continuation of the study of reflexive-insensitive logics (we borrow this terminology from [3]) that was initiated in [5] and further developed in [7] and [3]. In addition to this more systematic motive, it is also our desire to further elucidate (and exploit, in some cases) the connections between these logics and provability logics.<sup>1</sup>

---

\*We would like to thank the participants of the *Frontiers of Non-Classicality* conference for helpful conversations on the contents of this paper. We would also like to thank the organizers of that conference. We are grateful to the editors of this journal, and to the anonymous referee, whose insightful comments and suggestions improved the paper. This research was partially supported by FAPESP (São Paulo Research Foundation) grant number 2013/25095-4.

<sup>1</sup>In addition to the epistemic interpretations, reflexive-insensitive logics (RI-logics) also have a metaphysical interpretation, introduced by Marcos [5], where they are understood as formalizing the notion of a proposition being “accidentally true”. These logics are also

While finding applications throughout philosophical logic, the logics on which we focus actually have their genesis in paraconsistent logics. Specifically, they surfaced out of attempts to provide a modal interpretation of the consistency operator found in the logics of formal inconsistency (LFIs).<sup>2</sup> In the search for applications to relational structures of a metatheoretical operator of consistency (expressed in the object language with  $\circ$ ), it was investigated whether it is possible to modally interpret  $\circ$  in such a way that the resulting modal logics could still be considered LFIs (i.e., trivialization would occur when not only a formula  $\varphi$  and its negation were derivable, but also the formula  $\circ\varphi$ , expressing the consistency of  $\varphi$ ). Both a positive and a negative answer was found in [6]. On the one hand, it was shown that by interpreting  $\circ\varphi$  as  $\varphi \rightarrow \Box\varphi$ , and a paraconsistent negation  $\sim$  as  $\Diamond\neg$ , it is possible to obtain modal LFIs. On the other hand, any hope for discovering new interesting applications of LFIs by means of such an interpretation is unrealistic. The latter observation is expressed by the following result.

**Fact.** (Marcos [6]) *For every normal modal logic, a language  $\mathcal{L}$ , using  $\wedge$ ,  $\vee$ ,  $\rightarrow$ ,  $\neg$ ,  $\Box$ , and  $\Diamond$  as logical symbols, and a language  $\mathcal{L}'$ , using  $\wedge$ ,  $\vee$ ,  $\rightarrow$ ,  $\sim$ , and  $\circ$  as logical symbols, characterize the same logic under two different signatures.*

However, the above result can also be read as accentuating the non-classical character of normal modal logics, and this was indeed the reading proposed in [6]. Specifically, it concretely articulates the intuition that the boundary between the global, intensional, character of modal logic and the local, extensional, perspective is also part of the dividing line between non-classicality and classicality. Pursuing this investigation of the tenuous border between classical and non-classical, the study of logics using the above modal interpretation of the  $\circ$  operator and a classical negation gave rise to the definition and characterization of RI-logics, as in [3], by means of a relational semantics.

This paper, on the other hand, investigates the use of neighborhood frames for RI-logics, viewed epistemically as logics of unknown truths, and the related non-normal epistemic systems formalizing the notion of false belief. In the logic of unknown truths, first introduced by Steinsvold [8], instead of considering the more common epistemic modalities “is known” or “is believed”, the primitive modal operator is intended to formalize “is

---

the cousins of logics of contingency and non-contingency [4], for which a neighborhood approach was given recently in [11].

<sup>2</sup>See [1] for a general, and recent, presentation of these logics.

true, but not known”. In previous examinations of these logics, the semantics employed were either topological (in [8]), or relational (in [7] and [9]). To our knowledge, the appropriateness of neighborhood semantics for these logics has not yet been explored. The purpose of this paper is to develop just such a treatment.<sup>3</sup> In fact, we argue that neighborhood semantics may even be more appropriate because they accentuate properties of the logics that may go under-appreciated when examined through the lens of relational frames. For example, related to the expressivity of the logics, one can find appropriate neighborhood frames for the logic of unknown truths that do not correspond to any relational frames (whereas all appropriate relational frames have a pointwise-equivalent neighborhood frame). A suggestive conclusion one may draw from these observations is that, in some sense, the non-classicality of RI-logics resurfaces semantically in their weakness of expressivity, which is highlighted by the use of a finer-grained semantics.

In addition, it was initially hoped that switching to a neighborhood semantics might aid efforts to answer problems about the logics of false belief that were left open in [9], specifically issues relating to the appropriate W-logic of transitive frames. While relational frames thrive on the elegance of their binary accessibility relations, neighborhood semantics, while inarguably more encumbered with low-level set-theoretic machinery, allow for, in some cases, a greater extent of mathematical control over the frames, and thus their logics (hence allowing for the study of logics for which relational frames are not appropriate). And, in fact, the switch to neighborhood frames does shed some light on the behavior of the W-logics. Unfortunately, most of the results (in section 2.4, below) are negative, but they do suggest possible approaches to the open problems one might explore.

## 1 Logics of Unknown Truths

### 1.1 Language

Rather than the usual modal language where ‘ $\Box$ ’ or ‘ $\Diamond$ ’ is primitive, we will make use of a language,  $\mathcal{L}^\circ$ , in which ‘ $\circ$ ’ is the primitive modal operator. Let  $Var$  be a countable set of propositional variables. The members of  $Form_{\mathcal{L}^\circ}$ , the well-formed formulas of  $\mathcal{L}^\circ$ , are defined recursively as:

---

<sup>3</sup>In addition, given that neighborhood frames can be seen as a generalization of both relational and topological frames, it is our hope that studying the logics of unknown truths and false beliefs together, from the perspective of this more general framework, will unite the previous efforts.

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid \circ\varphi$$

For  $p \in Var$ . The usual definitions for  $\vee$ ,  $\rightarrow$  (material implication), etc., apply. In addition, we define  $\bullet$  so that  $\bullet\varphi$  stands for  $\neg\circ\varphi$ .

## 1.2 Axiomatization

In this paper, our basic logic will be the smallest set of formulas from  $Form_{\mathcal{L}^\circ}$  containing the following axioms, and closed under the given rules.

Axioms:

$PL$  All instances of propositional tautologies

$b0$   $\circ\top$

$b1$   $\bullet\varphi \rightarrow \varphi$

$b2$   $(\circ\varphi \wedge \circ\psi) \rightarrow \circ(\varphi \wedge \psi)$

Rules:

$MP$  Modus Ponens

$bN$  from  $\vdash \varphi \rightarrow \psi$  one can infer  $\vdash (\circ\varphi \wedge \varphi) \rightarrow (\circ\psi \wedge \psi)$

This logic, which we can call  $\mathbf{B}_K$ , is sound and complete with respect to the class of all relational frames. (This result was first proved in [5], and extended in [7]. A more comprehensive investigation into the soundness and completeness of these kinds of logics with respect to relational frames was conducted in [3].) The main purpose of this paper is to explore the adequacy of neighborhood semantics for  $\mathbf{B}_K$ , and potential extensions thereof.

## 1.3 Neighborhood Semantics

We briefly introduce the basic ideas surrounding neighborhood semantics in this section. All of the definitions are more or less standard. One should consult Chellas [2] for details.

**Definition 1.1** (Neighborhood Frame). *A neighborhood frame is a pair  $\langle W, n \rangle$  where  $W$  is a non-empty set of states and  $n : W \rightarrow \mathcal{P}(\mathcal{P}(W))$  is the neighborhood function ( $n(w)$  is the set of neighborhoods of  $w$ ). A neighborhood model is a triple  $\langle W, n, V \rangle$  where  $W$  and  $n$  are as before, and  $V : Var \rightarrow \mathcal{P}(W)$  is a valuation function.*

Truth at a state in a model for formulas can be defined recursively:

$$\begin{aligned} M, w \models p & \quad \text{iff} \quad w \in V(p) \\ M, w \models \neg\varphi & \quad \text{iff} \quad \text{it is not the case that } M, w \models \varphi \text{ (or } M, w \not\models \varphi) \\ M, w \models \varphi \wedge \psi & \quad \text{iff} \quad M, w \models \varphi \text{ and } M, w \models \psi \\ M, w \models \circ\varphi & \quad \text{iff} \quad \text{if } M, w \models \varphi \text{ then } \llbracket \varphi \rrbracket_M \in n(w) \end{aligned}$$

where  $\llbracket \varphi \rrbracket_M := \{w \in W \mid M, w \models \varphi\}$  (the *truth set* of  $\varphi$ ) (we will usually omit the subscript when the model is clear from the context). Though we are treating  $\circ$  as primitive, one can semantically understand  $\bullet$  as:

$$M, w \models \bullet\varphi \quad \text{iff} \quad M, w \models \varphi \text{ and } \llbracket \varphi \rrbracket_M \notin n(w)$$

It is for this reason that one might refer to these logics as the logics of unknown truths. As is often the case in neighborhood semantics, the presence of  $\varphi$ 's truth set in  $n(w)$  amounts to  $\varphi$  being known (or believed, when the focus is doxastic). As such,  $\bullet\varphi$  can be understood as saying that  $\varphi$  is true (at some world or state), but is not known.

A formula  $\varphi$  is said to be *globally true* in a model  $M$  just in case  $M, w \models \varphi$  for all  $w \in W$ . This will be denoted  $M \models \varphi$ . A formula is *valid* in a frame iff it is globally true in all models based on the frame. Lastly, a formula is *valid with respect to a class of frames* iff it is valid on each frame in the class.<sup>4</sup>

We use the following definitions in order to provide frame characterization theorems. Given a set  $W$ ,  $\mathcal{F} \subseteq \mathcal{P}(W)$  is said to

- be *supplemented* if for  $X, Y \in \mathcal{P}(W)$ ,  $X \in \mathcal{F}$  and  $X \subseteq Y$  implies  $Y \in \mathcal{F}$ ;
- be *closed under (binary) intersections* if for  $X, Y \in \mathcal{P}(W)$ ,  $X \in \mathcal{F}$  and  $Y \in \mathcal{F}$  implies  $(X \cap Y) \in \mathcal{F}$ ;
- *contain the unit* if  $W \in \mathcal{F}$ ;
- *contain the core* if  $\bigcap \mathcal{F} \in \mathcal{F}$ .
- *is consistent* if  $\emptyset \notin \mathcal{F}$  and  $\mathcal{F} \neq \emptyset$

$\mathcal{F}$  is said to be a *quasi-filter* if it is supplemented and closed under intersections. If it also contains the unit, then we simply call it a *filter*. If  $\mathcal{F}$  is supplemented and contains the core, then it is said to be *augmented*. (Hence, being augmented implies being a filter, but not vice versa.) This terminology can be extended to neighborhood frames in the obvious manner. For

<sup>4</sup>In the usual modal language, we have the following semantic clause for  $\Box$ :

$$M, w \models \Box\varphi \quad \text{iff} \quad \llbracket \varphi \rrbracket_M \in n(w)$$

Nothing else changes.

example, if  $n(w)$  is augmented for all  $w \in W$  in some frame, then we say that the frame itself is augmented.

Soundness of the axiomatization with respect to the semantics is readily forthcoming.

**Theorem 1.2.**  $\mathbf{B}_K$  is sound with respect to the class of filters.

*Proof.* The proof is straightforward, and not unusual. The case of  $b0$  makes use of the fact that filters contain the unit, closure under binary intersection is used for  $b2$ , and supplementation is needed for  $bN$ . ■

## 1.4 Completeness

In the following definition,  $|\varphi|_{\mathbf{S}}$  stands for the set of all maximal  $\mathbf{S}$ -consistent sets ( $\text{MaxCon}(\mathbf{S})$ ) that contain  $\varphi$ . Formally:

$$|\varphi|_{\mathbf{S}} := \{\Gamma \in \text{MaxCon}(\mathbf{S}) : \varphi \in \Gamma\}$$

We will often refer to  $|\varphi|_{\mathbf{S}}$  as the *proof set* of  $\varphi$ .

**Definition 1.3** (Canonical Model). *Let  $\mathbf{S}$  be a logical system extending  $\mathbf{B}_K$  and  $M = \langle W, n, V \rangle$  be a neighborhood model.  $M$  is said to be a canonical model for  $\mathbf{S}$  iff the following conditions hold:*

1.  $W = \{\Gamma : \Gamma \text{ is a maximal } \mathbf{S}\text{-consistent set of wff}\}$
2. For every  $\Gamma \in W$ ,  $(\circ\varphi \wedge \varphi) \in \Gamma$  iff  $|\varphi|_{\mathbf{S}} \in n(\Gamma)$
3.  $V(p) = |p|_{\mathbf{S}}$

When  $M$  is such a model, we will name it  $M_{\mathbf{S}} = \langle W_{\mathbf{S}}, n_{\mathbf{S}}, V_{\mathbf{S}} \rangle$ .

Note that  $n$  is well-defined. If  $|\varphi|_{\mathbf{S}} = |\psi|_{\mathbf{S}}$ , then  $\vdash_{\mathbf{S}} \varphi \leftrightarrow \psi$ . From the rule  $bN$ , we then have  $\vdash_{\mathbf{S}} (\circ\varphi \wedge \varphi) \leftrightarrow (\circ\psi \wedge \psi)$ . Therefore,  $(\circ\varphi \wedge \varphi) \in \Gamma$  iff  $(\circ\psi \wedge \psi) \in \Gamma$ , and  $|\varphi|_{\mathbf{S}} \in n_{\mathbf{S}}(\Gamma)$  iff  $|\psi|_{\mathbf{S}} \in n_{\mathbf{S}}(\Gamma)$ .

In addition, we will often refer to the *minimal* canonical model for a logic. This is the logic in which each  $n(w) = \{|\varphi|_{\mathbf{S}} : (\circ\varphi \wedge \varphi) \in w\}$ , and nothing more. In contrast, the maximal canonical model would result from adding, to this set, the set of all non-proof sets. Usually, we will only be concerned with the minimal canonical model.

(For the remainder of this section, we omit the  $\mathbf{S}$  subscript where doing so does not create problematic ambiguities.)

**Theorem 1.4.** *Let  $M$  be a canonical model for  $\mathbf{S}$ . Then we have that*

$$\llbracket \varphi \rrbracket_M = |\varphi|$$

*Proof.* We only show the cases for  $\circ$ .

$\Gamma \in \llbracket \circ\varphi \rrbracket$  iff either  $\Gamma \notin \llbracket \varphi \rrbracket$  or  $\llbracket \varphi \rrbracket \in n(\Gamma)$ .

If  $\Gamma \notin \llbracket \varphi \rrbracket$  then, by the induction hypothesis, we have that  $\Gamma \notin |\varphi|$ . This then immediately implies, from axiom  $b1$ , that  $\Gamma \in |\circ\varphi|$ .

If  $\llbracket \varphi \rrbracket \in n(\Gamma)$  then  $|\varphi| \in n(\Gamma)$ , which in turn implies, from the definition of the canonical model, that  $\circ\varphi \wedge \varphi \in \Gamma$ . Therefore,  $\Gamma \in |\circ\varphi|$ .

In the other direction, assume  $\Gamma \in |\circ\varphi|$ . If it is also the case that  $\Gamma \notin |\varphi|$ , then  $\Gamma \notin \llbracket \varphi \rrbracket$  and so  $\Gamma \in \llbracket \circ\varphi \rrbracket$ .

Otherwise, we have that  $\Gamma \in |\varphi|$  as well as  $\Gamma \in |\circ\varphi|$ . Thus  $\Gamma \in |\circ\varphi \wedge \varphi|$ . Then, by definition,  $|\varphi| \in n(\Gamma)$ , and so  $\llbracket \varphi \rrbracket \in n(\Gamma)$ . Therefore,  $\Gamma \in \llbracket \circ\varphi \rrbracket$ . ■

**Observation 1.5.** *The minimal canonical model for  $\mathbf{B}_K$  is not supplemented*

*Proof.* The most straightforward, though somewhat cumbersome, way to proceed is to pick an arbitrary  $X \subseteq W$  that is not a proof set (such an  $X$  must exist based on cardinality considerations), and consider, by cases, the relationship  $X$  could have with the sets  $|\varphi|$  and  $|\neg\varphi|$ , where both  $\varphi$  and  $\neg\varphi$  are consistent formulas. In every case, one can conclude that the model is not supplemented. ■

**Definition 1.6** (Supplementation). *Let  $M = \langle W, n \rangle$  be a neighborhood frame. Its supplementation,  $M^+ = \langle W, n^+ \rangle$ , is the structure in which  $n^+$  is defined such that for all  $w \in W$ :*

$$n^+(w) := \{Y \subseteq W : \exists X \in n(w) \text{ s.t. } X \subseteq Y\}$$

**Lemma 1.7.** *Let  $M$  be the smallest canonical model for  $\mathbf{B}_K$ . Take  $M^+ = \langle W, n^+, V \rangle$  to be its supplementation. Then  $M^+$  is also a canonical model for  $\mathbf{B}_K$ .*

*Proof.* We just have to show that the condition on canonical models continues to hold for  $n^+$ :

$$(\circ\varphi \wedge \varphi) \in \Gamma \text{ iff } |\varphi| \in n^+(\Gamma)$$

Left to right is immediate, since  $n(\Gamma) \subseteq n^+(\Gamma)$ . From right to left, assume that some  $|\varphi| \in n^+(\Gamma)$ . Then there must be some  $\psi$  for which  $|\psi| \subseteq |\varphi|$ , and  $(\circ\psi \wedge \psi) \in \Gamma$ . But if  $|\psi| \subseteq |\varphi|$ , then  $\psi \rightarrow \varphi$  is a theorem. Therefore,  $(\circ\psi \wedge \psi) \rightarrow (\circ\varphi \wedge \varphi)$  is a theorem, and so  $(\circ\varphi \wedge \varphi) \in \Gamma$ . ■

**Theorem 1.8.**  $\mathbf{B}_K$  is complete with respect to the class of filters.

*Proof.* The proof just consists in showing that there is a canonical model for  $\mathbf{B}_K$  that is supplemented, closed under binary intersection, and contains the unit. So just consider the supplementation of the minimal canonical model for  $\mathbf{B}_K$ ,  $M^+$ . It is obviously supplemented. Furthermore, from axiom  $b0$ ,  $\Gamma \in |\circ\top|$  for every  $\Gamma$ . Clearly, since  $\Gamma \in |\top|$  as well,  $W = |\top| \in n(\Gamma)$  for every  $\Gamma$ .

For binary intersections, assume  $X, Y \in n(\Gamma)$ . Then there must be a  $\varphi$  and  $\psi$  for which  $|\varphi| \subseteq X$  and  $|\psi| \subseteq Y$  and such that  $|\varphi| \in n(\Gamma)$  and  $|\psi| \in n(\Gamma)$ . Then  $(\circ\varphi \wedge \varphi) \in \Gamma$  and  $(\circ\psi \wedge \psi) \in \Gamma$ . Thus,  $(\varphi \wedge \psi) \in \Gamma$ , and, from axiom  $b2$ ,  $\circ(\varphi \wedge \psi) \in \Gamma$ . Therefore,  $|\varphi \wedge \psi| \in n(\Gamma)$ .  $|\varphi \wedge \psi| = |\varphi| \cap |\psi|$  and  $|\varphi| \cap |\psi| \subseteq X \cap Y$ , so  $X \cap Y \in n(\Gamma)$ , as desired. ■

It is also the case that we get another completeness result for free.

**Theorem 1.9.**  $\mathbf{B}_K$  is complete with respect to the class of augmented frames.

*Proof.* This is a direct consequence the completeness of  $\mathbf{B}_K$  with respect to relational frames, combined with the facts that  $\circ$  is definable in terms of  $\square$ , and every augmented model is pointwise equivalent to a relational model, and vice versa. ■

## 1.5 Mirror Reduction

In the context of relational semantics, when considering logics extending  $\mathbf{B}_K$  we encounter the phenomenon that two frames satisfy the same formulas of  $\mathcal{L}^\circ$  when they are mirror-related: when one frame can be obtained from the other simply through the addition and/or subtraction of reflexive arrows in the accessibility relation (a proof can be found in [5]).

We should be able to identify an analogous phenomenon in the context of neighborhood frames.<sup>5</sup>

<sup>5</sup>The guiding intuition is the following: if a formula  $\bullet\varphi$  holds at some state  $w$ , then we

Let  $M = \langle W, n, V \rangle$  be a model of  $\mathbf{BK}$ . For each  $w \in W$ , define  $S_w := \{X \in \mathcal{P}(W) : w \notin X\}$ . Furthermore let us define two new models  $M^{t^+} = \langle W, n^{t^+}, V \rangle$  and  $M^{t^-} = \langle W, n^{t^-}, V \rangle$  which are obtained from  $M$  by way of the following modifications of the neighborhood function:

$$n^{t^+}(w) := n(w) \cup S_w$$

and

$$n^{t^-}(w) := n(w) \setminus S_w$$

for each  $w \in W$ . ( $F^{t^+}$  and  $F^{t^-}$  are defined in the obvious manner.)

We then get the following theorem.

**Theorem 1.10.** *For all  $w \in W$  and  $\alpha \in \text{Form}_{\mathcal{L}^\circ}$ ,*

$$M, w \models \alpha \text{ iff } M^{t^+}, w \models \alpha$$

and

$$M, w \models \alpha \text{ iff } M^{t^-}, w \models \alpha$$

*Proof.* Both of these are proved by way of straightforward inductions on  $\alpha$ .

We will only detail the modal cases.

If  $M, w \models \circ\varphi$  then either  $M, w \not\models \varphi$  or  $\llbracket \varphi \rrbracket \in n(w)$ . In the first case, and in both inductions, we would then have that  $M^{t^+}, w \not\models \varphi$  and  $M^{t^-}, w \not\models \varphi$  by the induction hypotheses, which immediately give us that  $M^{t^+}, w \models \circ\varphi$  and  $M^{t^-}, w \models \circ\varphi$ .

Otherwise, we are in the situation where  $M, w \models \varphi$  (i.e.  $w \in \llbracket \varphi \rrbracket$ ) and  $\llbracket \varphi \rrbracket \in n(w)$ . There is nothing to show for  $M^{t^+}$ , since this model is constructed just by adding sets to each  $n(w)$ , and so the induction hypothesis immediately gives our result. For  $M^{t^-}$ , however, we point out that because  $\llbracket \varphi \rrbracket \notin S_w$ ,  $\llbracket \varphi \rrbracket$  will still be in  $n^{t^-}(w)$ . And so once again all we need is the induction hypothesis.

In the other direction, assume that  $M, w \not\models \circ\varphi$ . This then means that  $M, w \models \varphi$  and that  $\llbracket \varphi \rrbracket \notin n(w)$ . In this case it is the  $M^{t^-}$  that is immediate, since if  $\llbracket \varphi \rrbracket \notin n(w)$  then  $\llbracket \varphi \rrbracket \notin n^{t^-}(w)$ . In addition, because  $w \in \llbracket \varphi \rrbracket$ , we

---

should be able to add sets to  $n(w)$  so long as we are not adding the truth sets of any of the formulas that hold at  $w$ . On the other hand, to ensure that the truth of  $\circ\varphi$  is maintained, we must make sure not to remove the truth sets of true formulas. Thus, we can add and subtract sets at will, so long as they are not the truth sets of true formulas.

have that  $\llbracket \varphi \rrbracket \notin n(w)$  will imply that  $\llbracket \varphi \rrbracket \notin n^{t^+}(w)$ , as desired. Thus, the induction hypotheses finish the proof. ■

Notice that this theorem just talks about the endpoints of the phenomenon, when we add or subtract the sets  $S_w$  in their entirety. However, we needn't add or subtract the entire set  $S_w$  from the neighborhoods of  $w$ , but can add or subtract any subset of  $S_w$  without consequence.

**Corollary 1.11.** *If  $F$  is a frame for  $\mathbf{B}_K$ , then so is  $F^{t^+}$  and  $F^{t^-}$ , as well as the intermediate frames.*

We also point out that in the above canonical model construction we never have the situation in which  $w \notin \llbracket \varphi \rrbracket$  but  $\llbracket \varphi \rrbracket \in n(w)$  (in either the minimal canonical model or its supplementation). In other words, when  $M$  is the canonical model,  $M = M^{t^-}$ .

A few nice observations are now immediately forthcoming.

**Observation 1.12.** *No logic extending  $\mathbf{B}_K$  in the language  $\mathcal{L}^\circ$  can express the consistency of a frame.*

*Proof.* Since  $\circ\top$  is a theorem, it can never be the case that  $n(w) = \emptyset$ . But, because of the above result, we can always add  $\emptyset$  to any  $n(w)$  without consequence. ■

**Lemma 1.13.** *There are neighborhood frames of  $\mathbf{B}_K$  that are not closed under intersection and, therefore, are neither filters nor augmented.*

*Proof.* Consider the supplementation of the canonical frame, and notice that for a propositional variable  $p$ , the formula  $p \wedge \circ p$  is consistent, so will appear in some maximal set  $\Gamma$ . Therefore,  $\llbracket p \rrbracket \in n(\Gamma)$ . In addition, since it will never be the case that  $\perp \in \Gamma$ , we do not have  $\llbracket \perp \rrbracket \in n(\Gamma)$ . However, theorem 1.10 above tells us that we can harmlessly add the set  $\llbracket \neg p \rrbracket$  to  $n(\Gamma)$ . (Call the model that is the result of doing so  $M'$ ). Since  $\llbracket p \rrbracket \cap \llbracket \neg p \rrbracket = \emptyset \notin n(\Gamma)$ , we have violated the closure under intersections condition. Furthermore, this frame is a  $\mathbf{B}_K$  frame because it is indistinguishable, in  $\mathbf{B}_K$ , from the supplementation of the canonical frame.

(More simply, just take any model  $M$  in which, at some  $w$ ,  $M, w \models \varphi$  and  $\llbracket \varphi \rrbracket \in n(w)$ . Then  $\llbracket \neg \varphi \rrbracket \in S_w$ , and so is  $\emptyset$ . Thus, add  $\llbracket \neg \varphi \rrbracket$  to  $n$ , if it isn't already there, and remove  $\emptyset$ , if necessary.) ■

**Corollary 1.14.** *There are neighborhood frames of  $\mathbf{B}_K$  that do not have a pointwise-equivalent relational frame.*

*Proof.* Consider the model  $M'$  from the previous lemma. Notice that  $M', \Gamma \models \Box p \wedge \Box \neg p$ , but that it is not the case that  $M', \Gamma \models \Box \perp$ . However,  $(\Box p \wedge \Box \neg p) \rightarrow \Box \perp$  is a theorem of  $\mathbf{K}$ , and valid on all relational models. ■

Thus, there are models of  $\mathbf{B}_K$  that are not models for  $\mathbf{K}$ .

We can perform a similar construction that focuses on monotonicity (and supplementation) rather than intersections.

**Lemma 1.15.** *There are neighborhood frames of  $\mathbf{B}_K$  that are not supplemented.*

*Proof.* At any state  $w$  in the canonical model there will always be some  $\varphi$  for which  $\llbracket \varphi \rrbracket \notin n(w)$ . However, we can harmlessly add  $\llbracket \perp \rrbracket$  to  $n(w)$ . The resulting structure is not supplemented. In addition, on this new frame, considering the usual modal language, we will get that at  $w$  we have both  $\Box(\perp \rightarrow \varphi)$  and  $\Box \perp$ , but not  $\Box \varphi$ . ■

What this means in terms of expressivity is clear: the logic  $\mathbf{B}_K$  lacks the expressive power to ascertain pertinent semantic properties of its frames, leading to its being characterized both by the class of filters, as well as a distinct superclass.<sup>6</sup>

Semantically, it might suggest that, depending on one's aims, a neighborhood semantics might be a more interesting semantic tool for these logics: by restricting one's attention to the smaller class of frames, one obtains a more limited view of the semantic behavior of these logics.

## 1.6 Normal Extensions

A natural question concerns extensions of  $\mathbf{B}_K$ . This problem has been approached several times, in [5] and [7], and, most recently, in [3], where a translation between the usual modal language (which we can call  $\mathcal{L}^\Box$ , for convenience) and  $\mathcal{L}^\circ$  was used to provide reasonably robust techniques for importing characterization results. Specifically, consider the translation of formulas in  $\mathcal{L}^\Box$  to  $\mathcal{L}^\circ$  given by:

<sup>6</sup>Of course, this is not a particularly rare occurrence in modal logic, as characterization results for  $\mathbf{K}$  and  $\mathbf{S5}$ , for example, illustrate.

$$\begin{aligned}
p^\circ &= p \\
(\neg\alpha)^\circ &= \neg(\alpha^\circ) \\
(\alpha \wedge \beta)^\circ &= \alpha^\circ \wedge \beta^\circ \\
(\Box\alpha)^\circ &= \circ(\alpha^\circ) \wedge \alpha^\circ
\end{aligned}$$

We can extend this notation to sets of formulas, by stating that for a set of  $\mathcal{L}^\square$  formulas  $\Sigma$ ,  $\Sigma^\circ$  is just the set of  $\mathcal{L}^\circ$  formulas obtained from  $\Sigma$  by way of the above translation. With this, for example, one can show that if  $\mathbf{L}$  is a canonical normal modal logic axiomatized by some  $\Gamma$ , then, if its canonical model is contained in the class of (relational) frames  $\mathbb{C}$ ,  $\Gamma^\circ$  axiomatizes a logic (where the rules of inference are *MP* and *bN*) that is also complete with respect to  $\mathbb{C}$ . Call this logic  $\mathbf{L}^\circ$ .

The issue of soundness was also broached in [3], where a sufficient semantic condition for obtaining soundness results for  $\mathcal{L}^\circ$  from  $\mathbf{L}$  was given. It was shown that if the class of all  $\mathbf{L}$  frames, hereafter  $\mathbb{C}_{\mathbf{L}}$ , contains the reflexive closure of all of its frames, then  $\mathbf{L}^\circ$  is sound with respect to  $\mathbb{C}_{\mathbf{L}}$ . A corresponding syntactic condition, however, was not given.

We will improve upon this result in two ways. First, we show that the condition of containing all reflexive closures is in fact also a necessary condition. Second, we provide the desired sufficient syntactic condition as well, thereby providing a clearer understanding of the behavior of normal extensions of  $\mathbf{B}_K$ .<sup>7</sup>

Since we are now working with relational frames, the semantic condition for  $\circ$  is as follows:

$$M, w \models \circ\varphi \text{ iff either } M, w \not\models \varphi \text{ or, for all } x, wRx \text{ implies } M, x \models \varphi$$

In addition, we need to make use of a lemma connecting the satisfaction of formulas in  $\mathcal{L}^\circ$  with those in  $\mathcal{L}^\square$ .

**Lemma 1.16** ([3], Lemma 5.1). *Let  $M = \langle F, V \rangle$  be a model based on  $F = \langle W, R \rangle$  and take  $\alpha$  to be a formula of the language  $\mathcal{L}^\square$ . Then, for every  $x \in W$ :*

$$M, x \models \alpha^\circ \text{ iff } M^r, x \models \alpha$$

---

<sup>7</sup>Note that this is not irrelevant to the project of using neighborhood semantics to better understand these logics, given that, once a characterization of these logics is provided in either a relational or neighborhood setting, existing results on the correspondence between relational and neighborhood structures can be used to import this into the other domain, within which the domain specific semantic insensitivity results can be exploited.

where  $M^r$  is the model  $\langle F^r, V \rangle$  based on the frame  $F^r = \langle W, R^r \rangle$ , with  $R^r = R \cup \{ \langle x, x \rangle : x \in W \}$ .

**Lemma 1.17.** *Let  $\mathbf{L}$  be a normal modal logic. If  $\mathbf{L}^\circ$  is sound with respect to  $\mathbb{C}_{\mathbf{L}}$ , the class of all  $\mathbf{L}$ -frames, then  $\mathbb{C}_{\mathbf{L}}$  contains the reflexive closure of all of its frames.*

*Proof.* By way of contradiction, assume that  $\mathbf{L}^\circ$  is sound with respect to  $\mathbb{C}_{\mathbf{L}}$  but that  $\mathbb{C}_{\mathbf{L}}$  does not contain all reflexive closures. Then there is a frame  $F \in \mathbb{C}_{\mathbf{L}}$  such that  $F^r \notin \mathbb{C}_{\mathbf{L}}$  (i.e.  $F^r$  is not an  $\mathbf{L}$ -frame). Because  $F^r$  is not an  $\mathbf{L}$ -frame, we know that there is some theorem  $\varphi$  of  $\mathbf{L}$  that is not satisfied on some model  $M$  that is based on  $F^r$ . So  $F^r \not\models \varphi$ .

But since  $\varphi$  is a theorem of  $\mathbf{L}$ ,  $\varphi^\circ$  is a theorem of  $\mathbf{L}^\circ$ , and by the assumption of soundness, we have that  $\mathbb{C}_{\mathbf{L}} \models_{ri} \varphi^\circ$ . Obviously, then,  $F \models_{ri} \varphi^\circ$ . From lemma 1.16, we then get that  $F^r \models \varphi$ , a contradiction. ■

Thus, from the result already present in [3], one obtains the following theorem.

**Theorem 1.18.** *Let  $\mathbf{L}$  be a normal modal logic. Then the class of all  $\mathbf{L}$  frames,  $\mathbb{C}_{\mathbf{L}}$ , contains the reflexive closure of all of its frames if and only if  $\mathbf{L}^\circ$  is sound with respect to  $\mathbb{C}_{\mathbf{L}}$ .*

We can now give a syntactic condition on  $\mathbf{L}$  that can be proved to imply that  $\mathbb{C}_{\mathbf{L}}$  contains the reflexive closure of all its frames. To do so, we use the well-known boxdot-translation, which, as is shown below, is essentially the same as the  $\circ$ -translation given above.

**Definition 1.19.** *Let  $\Box\varphi$  be defined as  $\Box\varphi \wedge \varphi$ . Moreover let the boxdot-translation, from  $\mathcal{L}^\circ$  to  $\mathcal{L}^\Box$ , be given by*

$$\begin{aligned} p^\Box &= p \\ (\neg\alpha)^\Box &= \neg(\alpha^\Box) \\ (\alpha \wedge \beta)^\Box &= \alpha^\Box \wedge \beta^\Box \\ (\Box\alpha)^\Box &= \Box(\alpha^\Box) \wedge \alpha^\Box. \end{aligned}$$

First of all we notice that the operators  $\circ$  and  $\Box$  are inter-definable. Indeed, an easy calculation shows that  $\circ\varphi \wedge \varphi$  is equivalent to  $\Box\varphi$  and that  $\Box\varphi \vee \neg\varphi$  is equivalent to  $\circ\varphi$ .

Moreover, as remarked above, the  $\Box$ -translation is essentially equivalent to the  $\circ$ -translation. To be more precise, we have the following lemma<sup>8</sup>:

---

<sup>8</sup>[10] makes use of a similar connection between formulas in  $\mathcal{L}^\circ$  and formulas in  $\mathcal{L}^\Box$  to which the the boxdot-translation has been applied.

**Lemma 1.20.** *Let  $\alpha \in \mathcal{L}^\square$  by any modal formula. For any relational model  $M = \langle W, R, V \rangle$ , we have that*

$$M, w \models (\alpha)^\circ \text{ iff } M, w \models (\alpha)^\square$$

for any  $w \in W$ . Thus, for any frame

$$F \models (\alpha)^\circ \text{ iff } F \models (\alpha)^\square$$

*Proof.* This is demonstrated by a straightforward induction on  $\alpha$ . ■

**Theorem 1.21.** *Let  $\mathbf{L}$  be a normal modal logic.  $\mathbb{C}_{\mathbf{L}}$  contains the reflexive closure of all of its frames if*

$$\frac{\vdash \alpha}{\vdash (\alpha)^\square}$$

is an admissible rule of inference in  $\mathbf{L}$ .

*Proof.* We prove the contrapositive. If  $\mathbb{C}_{\mathbf{L}}$  does not contain the reflexive closure of all its frames, then there is some frame  $F \in \mathbb{C}_{\mathbf{L}}$  such that  $F^r \notin \mathbb{C}_{\mathbf{L}}$ . In other words, there must be some theorem of  $\mathbf{L}$ ,  $\alpha$ , such that  $\alpha$  is not valid on  $F^r$ . This, however, implies that  $\alpha^\circ$  is not valid on  $F$  (from lemma 1.16), and so, from lemma 1.20,  $\alpha^\square$  is not valid on  $F$ , and so cannot be a theorem of  $\mathbf{L}$ . Thus, the rule is not admissible. ■

Hence, we can see that the admissibility of this rule<sup>9</sup> amounts to a sufficient syntactic condition on the soundness of  $\mathbf{L}^\circ$  with respect to  $\mathbb{C}_{\mathbf{L}}$ . If we add an assumption about the completeness of  $\mathbf{L}$  with respect to  $\mathbb{C}_{\mathbf{L}}$ , then we can turn the conditional in the above theorem into a biconditional.

**Corollary 1.22.** *Let  $\mathbf{L}$  be a normal modal logic that is also complete with respect to  $\mathbb{C}_{\mathbf{L}}$ . Then  $\mathbb{C}_{\mathbf{L}}$  contains the reflexive closure of all of its frames if and only if*

$$\frac{\vdash \alpha}{\vdash (\alpha)^\square}$$

is an admissible rule of inference in  $\mathbf{L}$ .

<sup>9</sup>In [12], it is shown that in the presence of this rule,  $\mathbf{LT}$  (the logic axiomatized by adding  $\square\varphi \rightarrow \varphi$  to the axiomatization of  $\mathbf{L}$ ) is equivalent to  $\{\alpha : \mathbf{L} \vdash (\alpha)^\square\}$ .

## 2 Logics of false belief

### 2.1 Syntax, Axiomatization, and Semantics

We now turn our attention to the logics of false belief, introduced in [9]. Let  $\mathcal{L}^W$  be the language in which the well-formed formulas,  $Form_{\mathcal{L}^W}$  are defined by

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \varphi \mid W\varphi$$

We can follow the axiomatization provided in [9] (there referred to as  $S^W$ —we rename it in order to maintain more uniform naming conventions). Let  $\mathbf{A}_K$  be the smallest set of  $\mathcal{L}^W$  formulas containing the following axioms and closed under the rules.

Axioms:

- $PL$  All instances of propositional tautologies
- $a1$   $W\varphi \rightarrow \neg\varphi$
- $a2$   $(W\varphi \wedge W\psi) \rightarrow W(\varphi \wedge \psi)$

Rules:

- $MP$  Modus Ponens
- $aN$  from  $\vdash \varphi \rightarrow \psi$  one can infer  $\vdash (W\varphi \wedge \neg\psi) \rightarrow W\psi$

Semantically, the following clause is used for  $W$ :

$$M, w \models W\varphi \quad \text{iff} \quad M, w \not\models \varphi \text{ and } \llbracket \varphi \rrbracket_M \in n(w)$$

Given a doxastic understanding of neighborhoods, and the neighborhood function, this just says that  $\varphi$  is false, but is believed. Thus, it is a false belief.

In order to prove soundness and completeness results, we make use of the following definition:

**Definition 2.1** (Negatively Supplemented). *Let  $F = \langle W, n \rangle$  be a neighborhood frame. Call  $F$  negatively supplemented if (for all  $w \in W$  and  $X, Y \in \mathcal{P}(W)$ )  $X \in n(w)$ ,  $X \subseteq Y$ , and  $w \notin Y$  implies that  $Y \in n(w)$ .*

**Theorem 2.2.**  $\mathbf{A}_K$  is sound with respect to the class of all neighborhood frames that are closed under binary intersections and are negatively supplemented.

*Proof.* The validity of  $a1$  is immediate from the semantic definition of  $W$ .  $a2$  utilizes closure under intersections, and  $aN$  relies on negative supplementation. ■

## 2.2 Completeness

**Definition 2.3** ( $\mathbf{A_K}$  Canonical Model). *Let  $\mathbf{S}$  be a logical system extending  $\mathbf{A_K}$  and  $M = \langle W, n, V \rangle$  be a neighborhood model.  $M$  is said to be a canonical model for  $\mathbf{S}$  iff the following conditions hold:*

1.  $W = \{\Gamma : \Gamma \text{ is a maximal } \mathbf{S}\text{-consistent set of wff}\}$
2. For every  $\Gamma \in W$ ,  $\mathbb{W}\varphi \in \Gamma$  iff  $|\varphi|_{\mathbf{S}} \in n(\Gamma)$
3.  $V(p) = |p|_{\mathbf{S}}$

When  $M$  is such a model, we will name it  $M_{\mathbf{S}} = \langle W_{\mathbf{S}}, n_{\mathbf{S}}, V_{\mathbf{S}} \rangle$ .

The fact that the neighborhood function in canonical models is well defined follows from the fact that  $\vdash \varphi \leftrightarrow \psi$  implies that  $\vdash \mathbb{W}\varphi \leftrightarrow \mathbb{W}\psi$  (proved in [9]). Alternatively, assuming that  $|\varphi|_{\mathbf{S}} = |\psi|_{\mathbf{S}}$  and  $\mathbb{W}\varphi \in \Gamma$ , a1 gives that  $\neg\varphi \in \Gamma$ , and thus  $\neg\psi \in \Gamma$  as well. From aN,  $\vdash (\mathbb{W}\varphi \wedge \neg\psi) \rightarrow \mathbb{W}\psi$ , and so  $\mathbb{W}\psi \in \Gamma$ . An analogous argument works for  $\mathbb{W}\psi \in \Gamma$  implying  $\mathbb{W}\varphi \in \Gamma$ .

**Theorem 2.4.** *Let  $M$  be a canonical model for some  $\mathbf{S} \supseteq \mathbf{A_K}$ . Then*

$$\llbracket \varphi \rrbracket_M = |\varphi|$$

*Proof.* Assume  $\Gamma \in \llbracket \mathbb{W}\varphi \rrbracket$ . Then  $\Gamma \notin \llbracket \varphi \rrbracket$  and  $\llbracket \varphi \rrbracket \in n(\Gamma)$ . From the induction hypothesis,  $|\varphi| \in n(\Gamma)$ , and so, from the definition of the canonical model,  $\Gamma \in |\mathbb{W}\varphi|$ .

If  $\Gamma \in |\mathbb{W}\varphi|$ ,  $\Gamma \in |\neg\varphi|$  and  $|\varphi| \in n(\Gamma)$ . Then  $\Gamma \notin |\varphi|$  and so  $\Gamma \notin \llbracket \varphi \rrbracket$  but  $\llbracket \varphi \rrbracket \in n(\Gamma)$  (IH). From the semantic definition of  $\mathbb{W}$ ,  $\Gamma \in \llbracket \mathbb{W}\varphi \rrbracket$ . ■

**Definition 2.5** (Negative Supplementation). *Let  $M = \langle W, n \rangle$  be a neighborhood frame. Its negative supplementation,  $M^{\oplus} = \langle W, n^{\oplus} \rangle$ , is the structure in which  $n^{\oplus}$  is defined such that for all  $w \in W$ :*

$$n^{\oplus}(w) := \{Y \subseteq W : \exists X \in n(w) \text{ s.t. } X \subseteq Y \text{ and } w \notin Y\}$$

**Lemma 2.6.** *Let  $M$  be the smallest canonical model for  $\mathbf{A_K}$ . Take  $M^{\oplus}$  to be its negative supplementation. Then  $M^{\oplus}$  is also a canonical model for  $\mathbf{A_K}$ .*

*Proof.* As before, it only needs to be shown that the neighborhood function is maintained. That is:

$$\mathbf{W}\varphi \in \Gamma \text{ iff } |\varphi| \in n^\oplus(\Gamma)$$

Left to right is immediate, since  $n(\Gamma) \subseteq n^\oplus(\Gamma)$ .

In the other direction, assume that  $|\varphi| \in n^\oplus(\Gamma)$ . Then there is some  $\psi$  for which  $|\psi| \subseteq |\varphi|$  and  $\mathbf{W}\psi \in \Gamma$  and  $\Gamma \notin |\varphi|$ . Since  $|\psi| \subseteq |\varphi|$ ,  $\vdash (\mathbf{W}\psi \wedge \neg\varphi) \rightarrow \mathbf{W}\varphi$ . If  $\mathbf{W}\varphi \notin \Gamma$ , then either  $\mathbf{W}\psi \notin \Gamma$  or  $\neg\varphi \notin \Gamma$ . The first disjunct is immediately contradictory. So consider  $\neg\varphi \notin \Gamma$ . But this implies  $\varphi \in \Gamma$ , which contradicts  $\Gamma \notin |\varphi|$ . ■

**Corollary 2.7.**  $\mathbf{A}_K$  is complete with respect to the class of all neighborhood frames that are closed under binary intersections and are negatively supplemented.

*Proof.* The negative supplementation of the smallest canonical model is closed under intersections. ■

### 2.3 Frame Modifications

Before, when dealing with logics of unknown truths, we described modifications to the neighborhood function that did not affect the satisfiability of formulas. This corresponded to mirror reduction for relational frames. A similar question faces us now: what kinds of modifications can we make to the frames of the logics of false beliefs that do not affect the satisfaction of formulas? In fact, the relevant property here has a nice symmetry with our previous property.

Let  $M = \langle W, n, V \rangle$  be a model of  $\mathbf{A}_K$ . Consider the following set, for each  $w \in W$ :

$$U_w := \{X \in \mathcal{P}(W) : w \in X\}$$

Then, just like before, define two new models  $M^{u^+} = \langle W, n^{u^+}, V \rangle$  and  $M^{u^-} = \langle W, n^{u^-}, V \rangle$  where:

$$n^{u^+}(w) = n(w) \cup U_w$$

and

$$n^{u^-}(w) = n(w) \setminus U_w$$

(Similarly for  $F^{u^+}$  and  $F^{u^-}$ .)

**Theorem 2.8.** For all  $w \in W$  and  $\alpha \in \text{Form}_{\mathcal{L}^w}$ ,

$$M, w \models \alpha \text{ iff } M^{u^+}, w \models \alpha$$

and

$$M, w \models \alpha \text{ iff } M^{u^-}, w \models \alpha$$

*Proof.* Again, we only deal with the modal cases.

Assume  $M, w \models W\varphi$ . Then  $M, w \not\models \varphi$  and  $\llbracket \varphi \rrbracket \in n(w)$ . By the induction hypotheses,  $M^{u^+}, w \not\models \varphi$  and  $M^{u^-}, w \not\models \varphi$ . In the first case, for  $M^{u^+}$ , this is enough to give that  $M^{u^+}, w \models W\varphi$ , since  $\llbracket \varphi \rrbracket \in n(w)$  trivially implies  $\llbracket \varphi \rrbracket \in n^{u^+}(w)$ . For  $M^{u^-}$  the result follows from noticing that, since  $w \notin \llbracket \varphi \rrbracket$ ,  $\llbracket \varphi \rrbracket \notin U_w$ , and so  $\llbracket \varphi \rrbracket \in n^{u^-}(w)$  as well.

In the other direction, assume that  $M, w \not\models W\varphi$ . Then either  $M, w \models \varphi$  or  $\llbracket \varphi \rrbracket \notin n(w)$ .

If  $M, w \models \varphi$  then the induction hypotheses give that  $M^{u^+}, w \models \varphi$  and  $M^{u^-}, w \models \varphi$ . Thus,  $M^{u^+}, w \not\models W\varphi$  and  $M^{u^-}, w \not\models W\varphi$ .

So assume then that  $M, w \not\models \varphi$  but that  $\llbracket \varphi \rrbracket \notin n(w)$ . However, then  $w \notin \llbracket \varphi \rrbracket$ , and so  $\llbracket \varphi \rrbracket \notin U_w$ . Then, clearly,  $\llbracket \varphi \rrbracket$  is in neither  $n^{u^+}(w)$  nor  $n^{u^-}(w)$ . ■

**Corollary 2.9.** If  $F$  is a frame for  $\mathbf{A}_K$ , then so is  $F^{u^+}$  and  $F^{u^-}$ , as well as the intermediate frames.

**Observation 2.10.** No logic extending  $\mathbf{A}_K$  in the language  $\mathcal{L}^W$  can express that a frame contains the unit.

*Proof.* Since every  $w \in \llbracket \top \rrbracket$ , this set can always be removed from any  $n(w)$  without consequence. ■

## 2.4 The logic $\mathbf{K4}_W$

In [9], the problem of axiomatizing  $\mathbf{K4}_W$  (the logic of transitive relational frames in the language  $\mathcal{L}^W$ ) was posed, but so far little progress has been made to this end. In [9], it was already shown that the class of transitive frames could not be perfectly characterized in  $\mathcal{L}^W$  (due to the sorts of limitations of expressivity we have already encountered), but the question of characterizing the class of frames that is indiscernible (with respect to  $\mathcal{L}^W$ ) from the transitive class remains open. We believe that considering the neighborhood semantics for these logics can shed some light on the problem.

In the context of neighborhood semantics, transitivity (and validity of the the 4 axiom) corresponds to the following condition on frames, which we write as  $\tau(F)$  for a frame  $F = \langle W, N \rangle$ :

$$\forall w \in W (X \in n(w) \Rightarrow \{v \in W : X \in n(v)\} \in n(w)).$$

Abusing terminology, we call a neighborhood frame  $F$  transitive when  $\tau(F)$  holds.

**Theorem 2.11.** *No logic extending  $\mathbf{A}_K$  in the language  $\mathcal{L}^W$  can express that a frame is transitive.*

*Proof.* One can easily observe that for a transitive  $F$ , given a state  $w$  and set  $X \in n(w)$ ,  $w \in \{v : X \in n(v)\}$ . Therefore, by Theorem 2.8, we can always invalidate  $\tau(F)$ , for a frame  $F$ , without affecting the formulas in the language  $\mathcal{L}^W$ . ■

Therefore, no axiomatization can perfectly describe the class of transitive (neighborhood) frames.

In many ways, however, this is not at all surprising, given what we know about the expressive limitations of  $\mathcal{L}^W$  logics. In addition, a similar phenomenon is encountered with the  $\mathcal{L}^\circ$  logics, where their insensitivity to reflexivity always forces one to seek characterization results not for a particular class of relational or neighborhood frames, but rather the superclass incorporating, in addition, all those frames that are modally indiscernible to the ones in the initial class. This then poses the obvious question of ascertaining the appropriate superclass of the transitive neighborhood frames, and axiomatizing that class. However, this project also, and much more surprisingly, turns out to be problematic. In particular, no superclass (that is a proper subclass of the class of all frames) of the transitive neighborhood frames can be finitely axiomatized (in  $\mathcal{L}^W$ ).

**Definition 2.12.** *Given a frame  $F = \langle W, n \rangle$ , define its transitive closure  $F^{tc} = \langle W, n^{tc} \rangle$  inductively as  $\bigcup_{i \in \omega} F_i$ , with  $F_0 = F$  and  $F_{i+1} = \langle W, n_{i+1} \rangle$ , where  $n_{i+1} = tc(n_i)$ , defined as*

$$tc(n_i)(w) := n_i(w) \cup \{m_{n_i}(X) : X \in n_i(w)\}$$

for every  $w \in W$ , and

$$m_{n_i}(X) := \{z \in W : X \in n_i(z)\}$$

for  $X \subseteq W$ .

The transitive closure of a frame is a frame, since  $n^{tc}$  is a function (indeed, for every  $i \in \omega$ , we have  $n_i \subseteq n_{i+1}$ ), and it is transitive.

**Fact 2.13.** *If  $X \in n^{tc}(w) \setminus n(w)$ , then  $w \in X$ .*

*Proof.* Let  $i$  be the least natural number such that  $X \in n_{i+1}(w) \setminus n_i(w)$ . So  $X = \{z : Y \in n_i(z)\} = m_{n_i}(Y)$ , for some  $Y \in n_i(w)$ . Therefore  $w \in X$ . ■

**Theorem 2.14.** *Let  $M$  be a model based on a frame  $F$  and  $M^{tc}$  the corresponding one based on  $F^{tc}$ . For all  $w \in W$  and  $\alpha \in \text{Form}_{\mathcal{L}^W}$ , we have*

$$M, w \models \alpha \text{ iff } M^{tc}, w \models \alpha.$$

*Proof.* We only address the modal case. Assume  $M, w \models W\alpha$ . Then  $M, w \not\models \alpha$  and  $\llbracket \alpha \rrbracket \in n(w)$ . By the induction hypothesis,  $M^{tc}, w \not\models \alpha$ . Since  $\llbracket \alpha \rrbracket \in n(w) \subseteq n^{tc}(w)$ , we then have that  $M^{tc}, w \models W\alpha$ . In the other direction, assume that  $M, w \not\models W\alpha$ . Then either  $M, w \models \alpha$  or  $\llbracket \alpha \rrbracket \notin n(w)$ . If  $M, w \models \alpha$  then the induction hypothesis gives that  $M^{tc}, w \models \alpha$ . Therefore  $M^{tc}, w \not\models W\alpha$ . So assume  $M, w \not\models \alpha$ , but  $\llbracket \alpha \rrbracket \notin n(w)$ . Then  $w \notin \llbracket \alpha \rrbracket$  and so  $\llbracket \alpha \rrbracket \notin n^{tc}(w) \setminus n(w)$ . Thus,  $\llbracket \alpha \rrbracket \notin n^{tc}(w)$  and so, again,  $M^{tc}, w \not\models W\alpha$ . ■

**Theorem 2.15.** *Let  $\alpha$  be a consistent sentence in the language  $\mathcal{L}^W$ , and let  $\mathbb{C}_\alpha = \{F : F \models \alpha\}$ . If  $\mathbb{C}_\alpha$  is not the class of all frames, then there is a frame  $G$  such that  $G \notin \mathbb{C}_\alpha$  and  $\tau(G)$ .*

*Proof.* Let  $F = \langle W, n \rangle$  be a frame that does not belong to  $\mathbb{C}_\alpha$ . Then, by definition,  $F \not\models \alpha$ . If  $\tau(F)$  holds, then there is nothing to prove. If not then consider  $F^{tc} = \langle W, n^{tc} \rangle$ , the transitive closure of  $F$ . Then  $\tau(F^{tc})$ , but  $F^{tc} \not\models \alpha$ . ■

Theorem 2.15 represents an interesting application of the semantic insensitivity of the logics of false belief, that mirrors closely, but in a complimentary manner, the insensitivity to reflexivity of the logics of unknown truth<sup>10</sup>. Moreover, the fine-grained look that neighborhood semantics offers,

<sup>10</sup>The semantic parallel between the treatment of logics of unknown truths and of false beliefs suggests extending the methodology developed in [3] to the study of logics of false beliefs (using relational semantics) with the hope that such an approach might shed light on the open problems posed in [9]. Moreover, the fact that logics of unknown truths are insensitive to reflexivity, combined with their complementary semantic characterizations, suggests that one way of understanding the logics of false beliefs is as being insensitive to irreflexivity. However, the fact that the property of being irreflexive is not expressible in a

allows one to better appreciate the, in a certain sense, unstable definition transitivity receives in the language  $\mathcal{L}^W$ . However, the wider perspective that neighborhood semantics offers is not always instructive when compared with relational semantics. Indeed, in [9] it is shown that there exists a sentence that is valid in all transitive relational frames, but not in all relational frames. This fact does not contradict the above theorem, but tells us that the transitive closure of a neighborhood frame (like the one required for the proof of Theorem 2.15) may not necessarily be equivalent to any relational frame.

### 3 Conclusion

Although the closer look that neighborhood semantics offer on modal logic may make it harder to characterize specific classes of frames, nevertheless it may be used for discovering phenomena that might otherwise go unappreciated. For example, the nice complementary classifications that neighborhood semantics provide between the semantic insensitivity of the logics of false belief and those of unknown truths suggests exploring precisely the connection between the insensitivity to reflexivity of the RI-logics and the corresponding insensitivity (on relational semantics) of  $\mathcal{L}^W$  logics. This is planned for future work.

Coming back to the observations made in the introduction, this paper is the continuation of the study of the RI-logics started in [3]. As a matter of fact, one of the aims of this paper is to shed light on connections between seemingly separate areas of logic: epistemic logic and provability theory. Not only is the modal operator of the logics of unknown truths interdefinable with the  $\Box$ -operator, but the method of translation between normal modal logics and RI-logics, proposed in [3], is essentially the same as the so-called boxdot-translation. Therefore, a general study of the reflexive insensitive logics may shed light on issues in epistemic logic, formal metaphysics, as well as provability logic.

---

normal modal system poses an interesting obstacle to a straightforward transference of the translation-based technique of [3]. On the other hand, however, it does raise interesting questions about semantic methods for dealing with this insensitivity. We plan to develop this observation in future work, but for now it is worth stressing that, using neighborhood semantics, the semantic treatments of the logics of unknown truths and false beliefs mirror each other in a perfectly complementary fashion.

## References

- [1] W. Carnielli and M. Coniglio. *Paraconsistent Logic: Consistency, Contradiction and Negation*. Springer Verlag, 2016.
- [2] B. Chellas. *Modal Logic: an Introduction*. Cambridge University Press, 1980.
- [3] D. Gilbert and G. Venturi. Reflexive insensitive modal logics. *Review of Symbolic Logic*, 9(1):167–180, 2016.
- [4] L. Humberstone. The logic of non-contingency. *Notre Dame Journal of Formal Logic*, 36(2):214–229, 1995.
- [5] J. Marcos. Logics of essence and accident. *Bulletin of the section of logic*, 34(1):43–56, 2005.
- [6] J. Marcos. Nearly every normal modal logic is paranormal. *Logique et Analyse*, 48(189):279–300, 2005.
- [7] C. Steinsvold. Completeness for various logics of essence and accident. *Bulletin of the section of logic*, 37(2):93–101, 2008.
- [8] C. Steinsvold. A note on logics of ignorance and borders. *Notre Dame Journal of Formal Logic*, 49(4):385–392, 2008.
- [9] C. Steinsvold. Being wrong: Logics for false belief. *Notre Dame Journal of Formal Logic*, 52(3):245–253, 2011.
- [10] C. Steinsvold. The boxdot conjecture and the language of essence and accident. *Australasian Journal of Logic*, 10:18–35, 2011.
- [11] H. van Ditmarsch and J. Fan. Neighborhood contingency logic. In *Logic and its Applications*, Proceedings of the Sixth Indian Conference of the ICLA. Springer, 2015.
- [12] E. E. Zolin. Embeddings of propositional monomodal logics. *Logic Journal of the IGPL*, 8:861–882, 2000.