Semantics for Multi-Agent Only Knowing (extended abstract)

Arild Waaler[1,2] and Bjørnar Solhaug[3,4]

[1] Finnmark College, Norway
[2] Dep. of Informatics, University of Oslo, Norway
[3] SINTEF ICT, Norway
[4] Dep. of Information Science and Media Studies, University of Bergen, Norway

Abstract. The paper presents a Kripke semantics for a multi-agent generalization of Levesque's logic of "only knowing". We prove soundness and completeness and show that the logic has the Finite model property. The logic satisfies a Modal reduction theorem to the effect that any complex syntactical representation can be syntactically reduced to a provably equivalent form which directly reflects all the models of the representation. The model theory is illustrated by means of an example of supernormal defaults in a multi-agent language.

1 Introduction

Multi-agent belief logics can be viewed as systems designed for the representation of representations (or languages) that agents use for reasoning about other agents' cognitive states. A multi-agent *only knowing* system has language constructs for representing upper and lower bounds of beliefs; it thereby has constructs for expressing the exact content of an agent's belief state. The concept of only knowing has been studied by van der Hoek and Jaspars [4], who analyze the concept in relation to a number of epistemic systems. "Only knowing" based on K45 as the underlying system is, however, of particular interest since this system is capable of representing defeasible patterns of reasoning in a multi-agent context.

Generalizing the only knowing system of Levesque [6] to the multi-modal case is a non-trivial task; the tricky part of this is hidden in an axiom (which we shall refer to as the $\Diamond$-axiom) to the effect that $\Diamond\varphi$ (φ is logically possible) is an axiom *for each satisfiable, objective* φ ("objective" because it does not contain any modal operators). In a series of papers [5, 1–3] Halpern and Lakemeyer have attempted to formulate an appropriate generalization of this axiom in a multi-modal language; in the solution they end up with they enrich the object language with constructs for coding the satisfiability relation into the system. They also provide their analysis with a canonical model semantics. This semantics has, however, limited power, since the only model they allow is defined on the uncountable set of all maximally consistent sets. In particular, they do not have the Finite model property.

In [11] the first author introduced another generalization of the only knowing system of Levesque [6] to the multi-modal case which does not use meta-language operators like satisfiability. He proves consistency of the system L_I by proving that a complete subset of the language has a cut-free sequent calculus. He also proves that L_I is indeed equivalent to the system that Halpern and Lakemeyer claim is the correct multi-modal generalization of Levesque's system for the common part of the languages (i.e. formulae without the meta-concept operators).

The aim of this paper is to introduce a full-fledged Kripke semantics for L_I. No such results have previously been published. We first identify a kernel system of L_I called Æ_I; this system is a multi-modal generalization of the only knowing system analyzed in [7]. Æ_I has a natural Kripke semantics for which it is straightforward to show soundness, completeness and Finite model property. We show this in Sect. 2.

In Sect. 3 we identify L_I-models by restricting the class of $Æ_I$-models. The constructions used in proving completeness and Finite model property of $Æ_I$ carry over to L_I, although providing the appropriate generalizations of them is a laborious and non-trivial exercise. We prove completeness and soundness of L_I in Sect. 3.

In Sect. 4 we illustrate the power of the semantics by analyzing a representation of a multi-agent scenario in which an agent may have simple (i.e. supernormal) default assumptions about other agents' beliefs. By semantical means we establish the theoremhood of some illuminating equivalences which are not easily obtained by syntactic means.

In Sect. 5, we limit the language to finite languages, languages by means of which the Finite model property for L_I is established. A modal reduction property for L_I is addressed, which states that any "only knowing" expression is provably equivalent to a disjunction of "only knowing" expressions of a particular simple form. Each of these latter expressions provides us with an explicit syntactical representation of a particular model of the original formula. The reduction theorem hence provides us with a syntactical representation of all the models of a given "only knowing" formula. The latter expressions furthermore explicitly characterize the possible cognitive states of the agent, given the initial "only knowing" expression.

2 The Core System $Æ_I$

2.1 Syntax of $Æ_I$

The object language contains a stock of propositional letters, the Boolean connectives $\neg$ (negation), $\vee$ (disjunction), $\wedge$ (conjunction), $\supset$ (conditional), and $\equiv$ (biconditional), and modal operators B_a and C_a for each a in a non-empty index set I. The intended interpretation is that the index set I represents the set of agents, B_a is a belief operator, and C_a a complementary *co-belief* operator for agent a. We define the "all I know" modality O_a by $\mathsf{O}_a\varphi = \mathsf{B}_a\varphi \wedge \mathsf{C}_a\neg\varphi$. The function of the co-belief modality is to enable the representation of "believing a proposition φ at most," captured by the formula $\mathsf{C}_a\neg\varphi$. $\mathsf{O}_a\varphi$ thus expresses limits on the belief state of agent a from two directions, i.e. a believes at least that φ and a believes at most that φ. Further abbreviations: $\top$ is any tautology, $\bot$ any contradiction, b_a is $\neg\mathsf{B}_a\neg$, c_a is $\neg\mathsf{C}_a\neg$, $\Box_a\varphi$ is $\mathsf{B}_a\varphi \wedge \mathsf{C}_a\varphi$, the dual $\Diamond_a\varphi$ is $\mathsf{b}_a\varphi \vee \mathsf{c}_a\varphi$. The intended meaning of $\Box_a$ is logical necessity.

A formula φ is a *modal atom of modality* a if it is of the form $\mathsf{B}_a\psi$ or $\mathsf{C}_a\psi$ for an $a \in I$. φ is a *completely* a*-modalized* formula if it is a Boolean combination of modal atoms of modality a. If $J \subseteq I$, φ is *completely* J*-modalized* if it is a Boolean combination of modal atoms of modalities in J. φ is *free of modality* a if it is a Boolean combination of propositional letters and modal atoms not of modality a. φ is a *first-order formula* if, for each $a \in I$ and each subformula $\mathsf{B}_a\psi$ or $\mathsf{C}_a\psi$ in φ, ψ is free of modality a. When a formula φ is written $\varphi^{\setminus a}$, it means that it is free of a. If Γ is a set of formulae, $\Gamma^{\setminus a}$ denotes $\{\varphi \in \Gamma \mid \varphi$ free of modality $a\}$, and $\mathsf{Sf}(\Gamma)$ denotes the set of subformulae of the formulae in Γ. These conventions pertain in particular to languages: If $\mathcal{L}$ is a language, $\mathcal{L}^{\setminus a}$ denotes the set of formulae in $\mathcal{L}$ free of a.

The *modal depth* $\mathrm{m}(\varphi)$ of a formula φ expresses the nesting of modalities for alternating agents in φ. Its recursive definition is as follows. The modal depth of a purely Boolean formula is 0. Otherwise, if φ is $\mathsf{B}_a\psi$ or $\mathsf{C}_a\psi$, let Ψ be the set of modal atoms which occur as subformulae in ψ. Then $\mathrm{m}(\varphi)$ is the maximal number in $\{\mathrm{m}(\chi)+1 \mid \chi \in \Psi$ and χ is not a-modalized$\} \cup \{\mathrm{m}(\chi) \mid \chi \in \Psi$ and χ is a-modalized$\}$. Otherwise, the modal depth of φ is the maximal $\mathrm{m}(\psi)$ for a subformula ψ of φ. So if φ is purely Boolean, $\mathrm{m}(\mathsf{B}_a\mathsf{B}_a\varphi) = 1$, while $\mathrm{m}(\mathsf{B}_a\mathsf{B}_b\varphi) = 2$. That is, the function m is defined such that the prefixing of a formula of modality a with an a-modality does not increase the modal depth.

The deducibility relation '⊢' of the logic $Æ_I$ is defined as the least set that contains all tautologies, is closed under all instances of the rules

$$\frac{\vdash \varphi}{\vdash \Box_a \varphi}\ (\mathrm{RN}) \qquad \frac{\vdash \varphi \quad \vdash \varphi \supset \psi}{\vdash \psi}\ (\mathrm{MP})$$

and contains all instances of the following schemata:

K_{B}: $\mathsf{B}_a(\varphi \supset \psi) \supset (\mathsf{B}_a\varphi \supset \mathsf{B}_a\psi)$ $\quad$ K_{C}: $\mathsf{C}_a(\varphi \supset \psi) \supset (\mathsf{C}_a\varphi \supset \mathsf{C}_a\psi)$

$B_\Box$: $\mathsf{B}_a\varphi \supset \Box_a \mathsf{B}_a\varphi$ $\quad$ $\overline{B}_\Box$: $\neg\mathsf{B}_a\varphi \supset \Box_a\neg\mathsf{B}_a\varphi$

$C_\Box$: $\mathsf{C}_a\varphi \supset \Box_a \mathsf{C}_a\varphi$ $\quad$ $\overline{C}_\Box$: $\neg\mathsf{C}_a\varphi \supset \Box_a\neg\mathsf{C}_a\varphi$

T: $\Box_a\varphi \supset \varphi$

We write $\vdash \varphi$ if φ is theorem of $Æ_I$, and $\varphi_1, \ldots, \varphi_n \vdash \psi$ for $\vdash (\varphi_1 \wedge \cdots \wedge \varphi_n) \supset \psi$. $\Gamma \vdash \varphi$ means that there is a finite number of formulae $\gamma_1, \ldots, \gamma_n$ in Γ such that $\gamma_1, \ldots, \gamma_n \vdash \varphi$. If $\Gamma \vdash \bot$, Γ is *inconsistent*. Otherwise Γ is consistent, written $\mathrm{Con}(\Gamma)$. We will without reference use the well-known principles of modal logic, especially substitution of provable equivalents, the derived rule

$$\frac{\varphi_1, \ldots, \varphi_n \vdash \psi}{\mathsf{B}_a\varphi_1, \ldots, \mathsf{B}_a\varphi_n \vdash \mathsf{B}_a\psi}$$

and the corresponding rule for C_a. We shall frequently use the fact that any formula is provably equivalent to a first-order formula with the same modal depth (see Lemma 2 of [11]). It is easy to prove that $\Box_a$ is an S5 modality.

2.2 Semantics for $Æ_I$

A *frame* is a structure $(W, \{R_a, S_a \mid a \in I\})$, where W is a non-empty set of points and R_a and S_a are binary relations satisfying the following two conditions:

(f1) Let X be either R_a or S_a and Y be either R_a or S_a or their complements $\overline{R}_a$ or $\overline{S}_a$. Then the composition $X \circ Y \subseteq Y$.

(f2) $E_a = R_a \cup S_a$ is reflexive.

Note that in standard terminology two of the eight subconditions of $(f1)$ state that R_a and S_a are transitive, e.g. $R_a \circ R_a \subseteq R_a$, while two of them state that they are Euclidean, e.g. $R_a \circ \overline{R}_a \subseteq \overline{R}_a$.

Lemma 1. *E_a is an equivalence relation.*

Proof. Transitivity: $(R_a \cup S_a) \circ (R_a \cup S_a) \subseteq R_a \circ R_a \cup R_a \circ S_a \cup S_a \circ R_a \cup S_a \circ S_a$. By $(f1)$ each composition is included in $R_a \cup S_a$; hence $E_a \circ E_a \subseteq E_a$. Euclideanness: $(R_a \cup S_a) \circ \overline{(R_a \cup S_a)} \subseteq (R_a \circ \overline{R}_a \cap R_a \circ \overline{S}_a) \cup (S_a \circ \overline{R}_a \cap S_a \circ \overline{S}_a)$, which, by $(f1)$, is included in $\overline{R}_a \cap \overline{S}_a = \overline{R_a \cup S_a}$. □

An *a-cluster* is an equivalence class of W modulo E_a. Let C be an a-cluster. We define the *belief part* C^+ and the *co-belief part* C^- of C by: $C^+ = \{x \in C \mid xR_ax\}$ and $C^- = \{x \in C \mid xS_ax\}$. C is *bisected* if $C^+ \cap C^- = \emptyset$.

Lemma 2. $C = C^+ \cup C^-$.

Proof. Since $R_a \cup S_a$ is reflexive, either xR_ax or xS_ax. □

Lemma 3. $R_a \cap (C \times C) = C \times C^+$ *and* $S_a \cap (C \times C) = C \times C^-$.

Proof. Let $x, y \in C$. Assume first that xR_ay. By Euclideanness, yR_ay. This shows that $R_a \cap (C \times C) \subseteq C \times C^+$. Conversely, assume yR_ay. By $(f2)$, $x, y \in C$ implies that either xR_ay or xS_ay. In the latter case $S_a \circ R_a \subseteq R_a$ gives xR_ay; hence xR_ay in any case. S_a is treated symmetrically. □

An *Æ$_I$-model* $M = (W, \{R_a, S_a \mid a \in I\}, V)$ is a frame with a valuation function V, which maps each propositional letter onto a subset of W. The satisfiability relation $\vDash_x$, $x \in W$, is defined by

$$
\begin{aligned}
M \vDash_x p \quad & \text{iff} \;\; x \in V(p),\ p \text{ a propositional letter},\\
M \vDash_x \neg\varphi \quad & \text{iff} \;\; M \nvDash_x \varphi,\\
M \vDash_x \mathsf{B}_a\varphi \quad & \text{iff} \;\; \forall y\,(xR_ay \text{ implies } M \vDash_y \varphi),\\
M \vDash_x \mathsf{C}_a\varphi \quad & \text{iff} \;\; \forall y\,(xS_ay \text{ implies } M \vDash_y \varphi),
\end{aligned}
$$

and in the usual way for the other Boolean connectives. The *truth set* of ψ in M, $||\psi||$, is defined as $\{x \in W \mid M \vDash_x \psi\}$. We write $M \vDash_X \varphi$ iff $(\forall x \in X)(M \vDash_x \varphi)$. $\Gamma \vDash \varphi$ means that for all models, φ is true at all points which satisfy all formulae in Γ. We will in this section refer to an Æ$_I$-model simply as a *model*. Note that by Lemma 3, each point in C can see the same points through R_a and through S_a. Hence all points in an a-cluster C agree on every completely a-modalized formula in every model on the frame.

A formula is valid in a frame if it is true at all points in all Æ$_I$-models on the frame. If φ is valid in all frames, we write $\vDash \varphi$, and say that φ is *valid*.

We shall say that a logic has the *Finite model property* if its set of theorems is identical to the set of formulae valid in all models over finite frames.

Lemma 4. *Let C be an a-cluster, and let $\mathsf{O}_a\varphi$ be true at a point in C in a model on the frame. Then C is bisected.*

Proof. Assume that C is not bisected, i.e. that there is a point x in C such that both xR_ax and xS_ax. By assumption, $\mathsf{O}_a\varphi$ is true at x in the assumed model. Hence $\vDash_x \varphi$ and $\vDash_x \neg\varphi$. Contradiction. □

Theorem 1. *If* $\vdash \varphi$ *then* $\vDash \varphi$.

Proof. Axiom T is valid by reflexivity of E_a, and the axioms $B_\Box$, $C_\Box$, $\overline{B}_\Box$ and $\overline{C}_\Box$ are valid by corresponding conditions in $(f1)$. The other axioms and rules are standard. □

A set s of formulae is *maximal* if it is consistent, and every proper extension of it is inconsistent. We shall without proof use Lindenbaum's lemma: every consistent set has a maximal extension. Let W^c be the set of all maximal sets, $V^c(p) = \{s \in W^c \mid p \in s\}$, where p is a propositional letter, and let the binary relations R^c_a, S^c_a on W^c be defined by: sR^c_at iff for all φ, $\mathsf{B}_a\varphi \in s$ implies $\varphi \in t$, and sS^c_at iff for all φ, $\mathsf{C}_a\varphi \in s$ implies $\varphi \in t$. The *canonical model* M^c is defined as $(W^c, \{R^c_a, S^c_a \mid a \in I\}, V^c)$.

Since the modalities are normal, the *Truth Lemma* holds: $M^c \vDash_s \varphi$ iff $\varphi \in s$. Completeness can then be established in the usual way:

Theorem 2. *Æ$_I$ is strongly complete, i.e. if $\Gamma \vDash \varphi$ then $\Gamma \vdash \varphi$.*

Proof. It is sufficient to prove that M^c is a model, i.e we have to verify that M^c satisfies the frame conditions $(f1)$ and $(f2)$. Each subcondition of $(f1)$ is imposed on M^c by one of the axioms $B_\Box$, $C_\Box$, $\overline{B}_\Box$ and $\overline{C}_\Box$. $(f2)$ is imposed by axiom T. □

The filtration $M^\dagger = (W^\dagger, \{R_a^\dagger, S_a^\dagger \mid a \in I\}, V^\dagger)$ of M^c through Ψ, where Ψ is a set closed under subformulae, is defined as follows. $W^\dagger$ is the set of equivalence classes of W^c modulo $\sim_\Psi$, where the equivalence relation $\sim_\Psi$ on W^c is defined by: $s \sim_\Psi t$ iff $s \cap \Psi = t \cap \Psi$; the equivalence class of s modulo $\sim_\Psi$ is denoted $|s|$. $V^\dagger$ is a function which, for all propositional letters p in Ψ, satisfies: $V^\dagger(p) = \{|s| \mid p \in s\}$. Let us say that s and t *agree on a in Ψ* if, for all χ,

$$\mathsf{B}_a\chi \in s \cap \Psi \text{ iff } \mathsf{B}_a\chi \in t \cap \Psi \text{ and } \mathsf{C}_a\chi \in s \cap \Psi \text{ iff } \mathsf{C}_a\chi \in t \cap \Psi.$$

The binary relations $R_a^\dagger$ and $S_a^\dagger$ are then given by:

$$|s|R_a^\dagger|t| \text{ iff } \forall\chi(\mathsf{B}_a\chi \in s \cap \Psi \text{ implies } \chi \in t \cap \Psi) \text{ and } s \text{ and } t \text{ agree on } a \text{ in } \Psi,$$
$$|s|S_a^\dagger|t| \text{ iff } \forall\chi(\mathsf{C}_a\chi \in s \cap \Psi \text{ implies } \chi \in t \cap \Psi) \text{ and } s \text{ and } t \text{ agree on } a \text{ in } \Psi.$$

The *Filtration Theorem* holds for $Æ_I$:

Theorem 3. *For all $s \in W^c$ and $\varphi \in \Psi$, $M^\dagger \vDash_{|s|} \varphi$ iff $\varphi \in s$.*

Proof. By standard theory it is sufficient to show that $R_a^\dagger$ satisfies the following two conditions: (1) if $sR_a^c t$ then $|s|R_a^\dagger|t|$; (2) if $|s|R_a^\dagger|t|$ and $\mathsf{B}_a\varphi \in s \cap \Psi$, then $\varphi \in t \cap \Psi$, and that $S_a^\dagger$ satisfies the corresponding two conditions. $R_a^\dagger$ and $S_a^\dagger$ are clearly symmetrical, and condition (2) trivially holds by definition of $R_a^\dagger$. To prove that the first condition holds, assume that not $|s|R_a^\dagger|t|$. Then either there is a formula $\mathsf{B}_a\chi \in \Psi$ such that (i) $\mathsf{B}_a\chi \in s$ and $\chi \notin t$, (ii) $\mathsf{B}_a\chi \in s$ and $\mathsf{B}_a\chi \notin t$, or (iii) $\mathsf{B}_a\chi \notin s$ and $\mathsf{B}_a\chi \in t$, or there is a $\mathsf{C}_a\xi \in \Psi$ such that either (iv) $\mathsf{C}_a\xi \in s$ and $\mathsf{C}_a\xi \notin t$, or (v) $\mathsf{C}_a\xi \notin s$ and $\mathsf{C}_a\xi \in t$. In the first case, not $sR_a^c t$ by definition. In the second case, not $sR_a^c t$ by axiom $B_\Box$, in the third by $\overline{B}_\Box$, in the fourth by $C_\Box$, and in the fifth case, not sR_a^c by axiom $\overline{C}_\Box$. □

Lemma 5. *$M^\dagger$ is an $Æ_I$-model.*

Proof. All subconditions of the frame condition $(f1)$ are easily verified from the definition of $R_a^\dagger$ and $S_a^\dagger$; as to $(f2)$, assume that $R_a^\dagger \cup S_a^\dagger$ is not reflexive. Then there is an $s \in W^c$, and formulae $\mathsf{B}_a\chi$ and $\mathsf{C}_a\xi$ both in $s \cap \Psi$, such $\chi \notin s$ and $\xi \notin s$. But, by modal logic, both $\mathsf{B}_a(\chi \vee \xi)$ and $\mathsf{C}_a(\chi \vee \xi)$ must be in s. By axiom T, $(\chi \vee \xi) \in s$, i.e. either $\chi \in s$ or $\xi \in s$, and we have a contradiction. □

Theorem 4. *$Æ_I$ has the Finite model property.*

3 The Multi-Modal Only Knowing Logic $\mathbf{L}_I$

The system addressed in this section turns out to be a strict extension of $Æ_I$. Notational warning: Concepts which in the previous section were introduced for $Æ_I$ should in this section be understood relative to L_I, especially the provability relation $\vdash$, the notion of consistency and the canonical model M^c.

3.1 Syntax of $\mathbf{L}_I$

System L_I is defined as $Æ_I$ with axiom T replaced by the $\Diamond$-axiom:

$\Diamond_a\varphi$ is an axiom provided $\varphi \nvdash \bot$, φ free of modality a.

The $\Diamond$-axiom has a circular pattern, and an argument is needed in order to show that the circularity is not vicious. Note, incidentally, that consistency cannot be obtained by standard semantical methods. For to prove soundness of the $\Diamond$-axiom one needs to transform the consistency condition into a corresponding semantical condition in terms of satisfiability and this, in turn, requires a completeness argument. The standard way of proving completeness is via maximal consistent sets, but they can only be used if consistency has already been established.

Consistency of L_I has been established syntactically via a Cut-elimination theorem for a sequent calculus formulation of L_I [11]. It follows from the subformula property of cut-free proofs in the sequent calculus that the multi-modal extension of the syntactical form of the $\Diamond$-axiom is indeed well-defined.

By means of the sequent calculus the proof of the next lemma is easy; since the sequent calculus rules are outside the scope of this paper, the proof (which uses the $\Diamond$-axiom in an essential way) is left to the reader.

Lemma 6. *Let Γ be a set of completely J-modalized formulae and Δ be a set of formulae free of every modality in J. If Γ and Δ are both L_I-consistent, then $\Gamma \cup \Delta$ is also L_I-consistent.*

Lemma 7. *T is a theorem of L_I. Hence, L_I is an extension of $Æ_I$.*

Proof. We will show that the assumption that there is an instance of T that is not a theorem leads to a contradiction. Let $\Box_a\varphi \supset \varphi$ be an instance of T. We may without loss of generality assume that $\Box_a\varphi$ is first-order and hence that φ is free of modality a. By assumption, $\Box_a\varphi, \neg\varphi \nvdash \bot$, i.e. both $\Box_a\varphi$ and $\neg\varphi$ are L_I-consistent. The latter entails $\vdash \Diamond_a\neg\varphi$ by the $\Diamond$-axiom, i.e. $\vdash \neg\Box_a\varphi$. But this contradicts the consistency of $\Box_a\varphi$. □

3.2 Semantics of $\mathbf{L}_I$

In the following we will assume an $Æ_I$-model $M = (W, \{R_a, S_a \mid a \in I\}, V)$. Recall from Sect. 2.2 that for each agent a, E_a partitions the set of points W in a model into a-clusters. We can think of an a-cluster as a representation of the set of points that an agent a finds logically possible. In the semantics of L_I, two different a-clusters will always represent the same space of possibilities as far as the external world, including beliefs of other agents, is concerned; they differ only on the beliefs they ascribe to agent a.

We can think of the a-clusters as a partitioning of the set of points W along one dimension. Orthogonal to these lie the *$\overline{a}$-clusters*, defined as the equivalence classes induced by the relation $\bigcap_{b\in I\setminus\{a\}} E_b$. This is illustrated in Fig. 1, which depicts an a-cluster C and an $\overline{a}$-cluster D. Note that all points in D agree on the beliefs and co-beliefs of all agents except a.

The square representing $C \cap D$ consists of points which agree on the beliefs of all agents, i.e. points which are equivalent modulo $E_I = \bigcap_{b\in I} E_b$. Equivalence classes modulo E_I are called *I-clusters*. An I-cluster B is *Boolean saturated* if the following condition holds for every subset P of the propositional letters: There is a point $x \in B$ such that for each propositional letter p, $x \in V(p)$ iff $p \in P$. Informally, B is Boolean saturated if the points in B span the set of all propositional valuations. Hence, if B is Boolean saturated, then *for every propositionally satisfiable set Γ of purely Boolean formulae there is an $x \in B$ such that $\models_x \Gamma$.*

To define models for L_I we are interested in $Æ_I$-models that meet three saturation properties, the first of which is:

(s1) Each I-cluster is Boolean saturated.

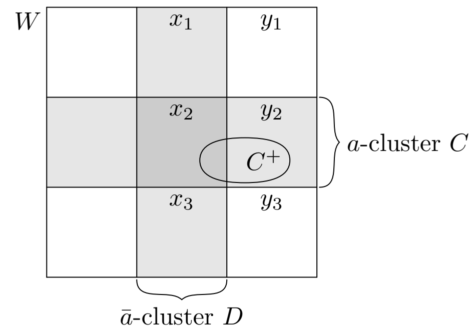

Fig. 1. A snapshot of a model, showing an a-cluster which intersects with an $\overline{a}$-cluster.

Let us say that a *cluster selection* is a set which consists of one and only one cluster for each $a \in I$. A cluster selection is *overlapping* if the intersection of the sets in it is non-empty. The second saturation property states that the model exhausts all possible ways of combining clusters:

(s2) Every cluster selection is overlapping.

The third condition employs notions which are easy to grasp from Fig. 1. Let us say that two points x and y have *equal Boolean force*, written $x \approx y$, if and only if they agree on propositional letters, i.e. for each propositional letter p, $x \in V(p)$ iff $y \in V(p)$. Note that points are related in this way across the model, irrespective of their cluster memberships. Let us e.g. assume that the six x and y points in Fig. 1 have equal Boolean force.

Two points x and y in W are *a-similar*, written $x \approx_a y$, if they are in the same $\overline{a}$-cluster and have equal Boolean force. Hence, if $x \approx_a y$, then x and y agree on all formulae free of modality a. The equivalence class of x modulo $\approx_a$ is denoted $[x]_a$. If $U \subseteq W$, let $[U]_a = \{[x]_a \mid x \in U\}$. In Fig. 1 $[x_1]_a = [x_2]_a = [x_3]_a$ and $[y_1]_a = [y_2]_a = [y_3]_a$. The two conditions defined so far entail that every a-cluster span the whole set of possible states of affairs, given the language of agent a:

Lemma 8. *Let C be any a-cluster in an $Æ_I$-model which satisfies (s1) and (s2). Then $[W]_a = [C]_a$.*

Proof. Clearly, $\{[x]_a \mid x \in C\} \subseteq [W]_a$. Conversely, assume that $[x']_a \in [W]_a$ and $x' \notin C$. Assume that x' belongs to the a-cluster C' and let S' be any cluster selection which contains C'. Let S be S' with C' replaced by S. By (s2), the I-cluster $D = \cap S \neq \emptyset$. Let $D' = \cap S'$; note that $x' \in D'$. By (s1) there is an $x \in D$ such that $x \approx x'$. We must then have that $x \approx_a x'$, i.e. $[x]_a \in [W]_a$. The last statement in the lemma follows immediately from definition. □

Let us say that a subset X of $[W]_a$ is *definable* if there is a set Γ of formulae free of modality a such that for each point x, $M \vDash_{[x]_a} \Gamma$ iff $[x]_a \in [W]_a$. $X \subseteq [W]_a$ is *N-definable* if the same condition holds, but with the proviso that the modal depth of each formula in Γ is at most N.

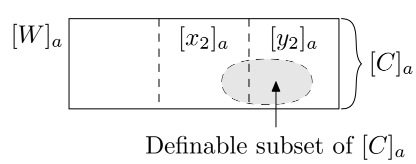

Fig. 2. The abstraction over W where each point is mapped to the set of all its a-similar points.

$[W]_a$ is an abstraction over W defined by means of a-clusters. Fig. 2 illustrates $[W]_a$ constructed over the W in Fig. 1. Note that for each point x and each set Γ of formulae free of modality a, $M \vDash_x \Gamma$ iff $M \vDash_{[x]_a} \Gamma$ (recall that $M \vDash_X \varphi$ iff $(\forall x \in X)(M \vDash_x \varphi)$). We can hence without ambiguity talk about a formula free of modality a being true in $[x]_a$ in $[W]_a$.

The belief state C^+ of the a-cluster C is depicted in Fig. 1; assume that there is a formula φ free of a such that φ is true at all points in C^+ and false at other points in C. The corresponding set $[C^+]_a$, depicted in Fig. 2, is then defined by φ. $[C^+]_a$ is, more precisely, m(φ)-defined by φ.

The third saturation property reflects the intuitive meaning of the $\Diamond$-axiom in the form of properties shared by all a-clusters. The condition can be conveniently expressed in terms of the set $[W]_a$.

(s3) For each definable $X \subseteq [W]_a$ and $Y \subseteq [W]_a$ such that $X \cup Y = [W]_a$ there is an a-cluster C such that $X = [C^+]_a$ and $Y = [C^-]_a$.

An $Æ_I$-model is *saturated* if it satisfies conditions (s1)-(s3). It is *N-saturated* if the clause 'definable' in (s3) is replaced by 'N-definable'. A formula is *valid* if it is true in all saturated models. It is *satisfiable* if it is true at some point in some saturated models and *N-satisfiable* if it holds in some N-saturated model.

Proposition 1. *Let Γ be a set of formulae whose modal depths are all less than a fixed integer N and $M = (W, \{R_a, S_a \mid a \in I\}, V)$ be any N-saturated model. Then Γ is satisfiable iff it is satisfied in M.*

Proof. "If" holds since N-satisfiability trivially entails satisfiability. Conversely, let $I = \{a_1, a_2, \dots\}$ and assume that Γ is satisfiable. First, observe that there are sets $\Gamma_0, \Gamma_1, \Gamma_2, \dots$ such that $\cup_{i\geq 0}\Gamma_i$ is satisfiable, where Γ_0 consists of purely Boolean formulae and Γ_i consists of formulae with modal depth $\leq N$ of one of the following forms: $\mathsf{B}_{a_i}\varphi$, $\mathsf{b}_{a_i}\varphi$, $\mathsf{C}_{a_i}\varphi$ or $\mathsf{c}_{a_i}\varphi$. To see this, transform each formula in Γ to a first-order formula and then use standard propositional reasoning.

Let $\Gamma_i^+ = \{\varphi \mid \mathsf{B}_{a_i}\varphi \in \Gamma\} \cup \{\top\}$ and $\Gamma_i^- = \{\varphi \mid \mathsf{C}_{a_i}\varphi \in \Gamma\} \cup \{\top\}$. Let X be the set of all $[z]_{a_i}$ which satisfies Γ_i^+ and Y be the set of all $[z]_{a_i}$ which satisfies Γ_i^-. Γ_i^+ then defines X and Γ_i^- defines Y.

First claim: $X \cup Y = [W]_{a_i}$. To see this, assume the contrary. Then there are formulae $\varphi \in \Gamma_i^+$ and $\psi \in \Gamma_i^+$ both free of modality a_i, and a point z such that $\neg(\varphi \vee \psi)$ is true at $[z]_{a_i}$. By assumption Γ is satisfiable; assume that M' satisfies Γ at u. By induction hypothesis M' satisfies $\neg(\varphi \vee \psi)$. By Lemma 8, there is a point v in the same a-cluster as u which satisfies $\neg(\varphi \vee \psi)$. But then $\mathsf{B}_{a_i}\varphi$ and $\mathsf{C}_{a_i}\psi$ cannot both be true at u, contradicting the assumption that u satisfies Γ. This proves the first claim.

By (s3), there is an a_i-cluster C_i in M such that $X = \{[x]_{a_i} \mid x \in C_i^+\}$ and $Y = \{[x]_{a_i} \mid x \in C_i^-\}$. Second claim: $M \vDash_{C_i} \Gamma_i$. By construction it is clear that each formula $\mathsf{B}_{a_i}\varphi$ in Γ is true in C_i. Let $\mathsf{b}_{a_i}\psi \in \Gamma$. By truth definition, $\Gamma_i^+ \cup \{\psi\}$ is satisfiable. By induction hypothesis this set is true at a point in M, say z. By Lemma 8 and construction, $[z]_{a_i} \in X$ and hence, $M \vDash_{C_i} \mathsf{b}_{a_i}\psi$. The same argument applies to formulae $\mathsf{C}_{a_i}\varphi$ and $\mathsf{c}_{a_i}\psi$ in Γ. This proves the second claim.

Let $C = \cap_{i\in I}C_i$. By (s2) C is an I-cluster. By (s1) there is a $z \in C$ which satisfies Γ_0. By construction, $M \vDash_z \cup_{i\geq 0}\Gamma_i$. □

Theorem 5. *Let φ be free of a. If φ is satisfiable, then $\Diamond_a\varphi$ is valid.*

Proof. Let C be any a-cluster in any saturated model M. By Proposition 1 there is an x such that $M \vDash_x \varphi$. By Lemma 8, there is a point $y \in C$ such that $x \approx_a y$, i.e. φ holds at a point in C. Thus $M \vDash_C \Diamond_a\varphi$. □

3.3 Completeness and Soundness

Let M^c be the canonical model for L_I. Since L_I is an extension of Æ_I it follows immediately that it is an Æ_I-model.

Lemma 9. *M^c is a saturated Æ_I-model.*

Proof. It is straightforward to show that for all maximal sets s and t: (i) s and t are in the same J-cluster iff s and t agree on all J-modalized formulae; (ii) $s \approx_a t$ iff s and t agree on all formulae free of modality a.

(s1): Let C be any I-cluster in W^c. Let Γ be $\cap C$. Γ contains the set of all completely modalized formulae true at all points in C. Let P be any set of propositional letters and $\overline{P}$ be the set of propositional letters not in P. By Lemma 6 $P \cup \overline{P} \cup \Gamma$ is consistent. By (i) and Lindenbaum's lemma it is contained in a point in C.

(s2): Let S be a cluster selection and $S^{\cap} = \{\cap C \mid C \in S\}$. By Lemma 6, $\cup S^{\cap}$ is consistent. By Lindenbaum's lemma it is contained in a maximal set, which proves (s2).

(s3): Let X and Y be definable subsets of $[W^c]_a$ such that $X \cup Y = [W^c]_a$; let Ξ define X and Υ define Y. We first show that (iii) if φ and ψ are free of a such that $\Xi \vdash \varphi$ and $\Xi \vdash \psi$, then $\vdash \varphi \vee \psi$. By (ii), $\varphi \in s$ for each s such that $[s]_a \in X$ and $\psi \in s$ for each s such that $[s]_a \in Y$. By Lemma 8, for each s, either $[s]_a \in X$ or $[s]_a \in Y$. Hence $\varphi \vee \psi$ is a member of every maximal set, and $\vdash \varphi \vee \psi$ holds by Lindenbaum's lemma.

Let $\Gamma = \{\mathsf{B}_a\varphi^{\setminus a} \mid \Xi \vdash \varphi^{\setminus a}\} \cup \{\mathsf{b}_a\varphi^{\setminus a} \mid \Xi \cup \{\varphi\}$ is L_I-consistent$\}$ and $\Delta = \{\mathsf{C}_a\varphi^{\setminus a} \mid \Upsilon \vdash \varphi^{\setminus a}\} \cup \{\mathsf{c}_a\varphi^{\setminus a} \mid \Upsilon \cup \{\varphi\}$ is L_I-consistent$\}$. We prove that (iv) $\Gamma \cup \Delta$ is L_I-consistent. Assume not; then there are formulae $\gamma, \delta, \chi_1, \ldots, \chi_m, \theta_1, \ldots, \theta_n$ free of a such that $\mathsf{B}_a\gamma, \mathsf{C}_a\delta, \mathsf{b}_a\chi_1, \ldots, \mathsf{b}_a\chi_m, \mathsf{c}_a\theta_1, \ldots, \mathsf{c}_a\theta_n \vdash \bot$. $\{\mathsf{B}_a\gamma, \mathsf{C}_a\delta\}$ is consistent, otherwise there are φ and ψ are free of a such that $\Xi \vdash \varphi$ and $\Xi \vdash \psi$ and $\mathsf{B}_a\varphi, \mathsf{C}_a\psi \vdash \bot$. By the $\Diamond$-axiom, $\nvdash \varphi \vee \psi$, contradicting (iii). Should $\Gamma \cup \Delta$ still be inconsistent, there must be a χ_i such that $\mathsf{B}_a\gamma, \mathsf{C}_a\delta \vdash \mathsf{B}_a\neg\chi_i$ or a θ_j such that $\mathsf{B}_a\gamma, \mathsf{C}_a\delta \vdash \mathsf{C}_a\neg\theta_j$. But by construction of Γ and Δ this is impossible.

Thus $\Gamma \cup \Delta$ is consistent. Let $\Gamma \cup \Delta \subseteq u$ and C be the a-cluster which u belongs to. To establish (s3) it is sufficient to show that for each $s \in C$, $s \in C^+$ iff $[s]_a \in X$ and $s \in C^-$ iff $[s]_a \in Y$. We prove the first. By definition and Truth lemma, $[s]_a \in X$ iff $\Xi \subseteq s$; we hence show that $s \in C^+$ iff $\Xi \subseteq s$. So let $s \in C^+$. By definition, $\{\varphi^{\setminus a} \mid \Xi \vdash \varphi^{\setminus a}\} \subseteq s$; in particular $\Xi \subseteq s$. Conversely, assume that $s \in C \setminus C^+$. By (i), there is a $\varphi^{\setminus a} \in s$ such that $\mathsf{b}_a\varphi^{\setminus a} \notin u$. By construction of Γ, $\Xi \cup \{\varphi^{\setminus a}\}$ is consistent, i.e. $\Xi \nsubseteq s$. This completes the proof. □

Theorem 6. *L_I is sound and strongly complete.*

Proof. Completeness follows directly from Lemma 9. All rules except the $\Diamond$-axiom are valid for the same reason as for Æ_I (cnf. Lemma 1). To prove that the $\Diamond$-axiom is also valid, assume that $\varphi^{\setminus a}$ is L_I-consistent. By Truth lemma, $\varphi^{\setminus a}$ is true at a point s in M^c. By Lemma 9, $\varphi^{\setminus a}$ is satisfiable. Conclude by Proposition 1. □

Theorem 7. *A formula φ is satisfiable iff it is $m(\varphi)$-satisfiable.*

Proof. Assume that φ is satisfiable. By soundness, φ is consistent. It is hence satisfied in M^c, which by Lemma 9 is $\mathrm{m}(\varphi)$-saturated. □

4 Example: Belief Representations with Defaults

We will in this section analyze the representation of supernormal defaults in the multi-modal language. A default for agent a will be represented as a conditional of the form $\mathsf{b}_a\psi \supset \psi$ within the scope of an O_a-modality. The default conditional states that agent a believes ψ in case ψ is compatible with the rest of the beliefs. We will use κ for the rest of the beliefs and assume that both κ and ψ are free of modality a.

The analysis amounts to breaking down a representation $\mathsf{O}_a\varphi$ into a disjunction of formulae of the form $\mathsf{O}_a\varphi'$, where φ' is free of modality a. This is a special case of the Modal Reduction Theorem addressed in Sect. 5.2.

To prove equivalences of the form $\vdash \mathsf{O}_a\varphi \equiv \mathsf{O}_a\varphi'$ we may argue semantically by showing that $M \models_C \mathsf{O}_a\varphi$ iff $M \models_C \mathsf{O}_a\varphi'$ for every a-cluster C in a model M; provable equivalence then follows by completeness. To this end, observe that if $M \models_x \varphi \equiv \varphi'$ for every x in an a-cluster C satisfying $\mathsf{O}_a\varphi$, then $M \models_C \mathsf{O}_a\varphi'$. The following observation is also useful.

Lemma 10. *Let C be an a-cluster in a saturated $Æ_I$-model M, and let ψ be a formula free of modality a. If ψ is satisfiable, there is a point $x \in C$ such that $M \models_x \psi$.*

Proof. By Proposition 1, ψ is satisfiable in any $\mathrm{m}(\psi)$-saturated model. Since M is saturated, M is also $\mathrm{m}(\psi)$-saturated, and so ψ is satisfied in M. Let D be an $\overline{a}$-cluster in M satisfying ψ. By (s2), $C \cap D \neq \emptyset$. Since all points in D agree on all completely b-modalized formulae, $a \neq b$, and since $C \cap D$ by (s1) is Boolean saturated, there is a point $x \in C \cap D$ such that $M \models_x \psi$. Hence ψ is satisfied in C. □

Example 1. We will first analyze the only knowing formula $\mathsf{O}_a\varphi_1$, where $\varphi_1 = \kappa \wedge (\mathsf{b}_a\psi \supset \psi)$. The content of $\mathsf{O}_a\varphi_1$ is made explicit by means of an equivalence which depends on the relationship between κ and ψ. There are two cases, listed in the left column below. We shall in both cases argue for the corresponding provable equivalence to the right.

1. $\kappa \vdash \neg\psi$ $\quad\quad$ $\vdash \mathsf{O}_a\varphi_1 \equiv \mathsf{O}_a\kappa$
2. $\mathrm{Con}(\kappa, \psi)$ $\quad\quad$ $\vdash \mathsf{O}_a\varphi_1 \equiv \mathsf{O}_a(\kappa \wedge \psi)$

Case 1. By soundness, $\kappa \models \neg\psi$. Let $M \models_C \mathsf{O}_a\varphi_1$. Then $C^+ \subseteq ||\kappa|| \subseteq ||\neg\psi||$, and hence (i) $M \models_C \mathsf{B}_a\neg\psi$. Let x be any point in C. If $M \models_x \varphi_1$, $M \models_x \kappa$ follows immediately. Conversely, suppose that $M \models_x \kappa$. By (i), $M \models_x \kappa \wedge \mathsf{B}_a\neg\psi$, and $M \models_x \varphi_1$ follows by PL. This proves that $M \models_x \varphi_1 \equiv \kappa$ for every $x \in C$, i.e. that $\mathsf{O}_a\varphi_1 \models \mathsf{O}_a\kappa$. For the other direction, assume that $M \models_C \mathsf{O}_a\kappa$. Then $C^+ = ||\kappa|| \subseteq ||\neg\psi||$, and statement (i) holds. The rest of the argument is as for the first direction.

Case 2. Assume first that $M \models_C \mathsf{O}_a\varphi_1$. Then (ii) $C^- \subseteq ||\kappa \supset \neg\psi||$. Since $\mathrm{Con}(\kappa, \psi)$, completeness entails that $\kappa \wedge \psi$ is satisfiable. By Lemma 10, there is a point $x \in C$ such that $M \models_x \kappa \wedge \psi$. By (ii), $x \in C^+$, and hence (iii) $M \models_C \mathsf{b}_a\psi$. Let $x \in C$ be any point such that $M \models_x \varphi_1$. By (iii), $M \models_x \varphi_1 \wedge \mathsf{b}_a\psi$, and so $M \models_x \kappa \wedge \psi$ by PL. Conversely, let $M \models_x \kappa \wedge \psi$; $x \in C$. By PL, we immediately have $M \models_x \varphi_1$. For the other direction, assume that $M \models_C \mathsf{O}_a(\kappa \wedge \psi)$. Then $C^- = ||\kappa \supset \neg\psi||$. Statement (iii) then holds, and the rest of the argument is as for the first direction.

In [3], Halpern and Lakemeyer show in two examples that their logic can be used for nonmonotonic reasoning. The belief representations are only knowing formulae involving one supernormal default. We will in the following two examples show that the logic L_I can be used to carry out the same derivations.

Example 2. Let p be agent a's secret and suppose he makes the assumption that unless he believes that b believes his secret, he assumes that she does not believe it. We will now prove that if this and κ is all he knows and if κ does not imply that b believes his secret, then he believes that she does not believe it. Formally, we show

$$\mathsf{O}_a(\kappa \wedge (\mathsf{b}_a\neg\mathsf{B}_b p \supset \neg\mathsf{B}_b p)) \vdash \mathsf{B}_a\neg\mathsf{B}_b p.$$

Our belief representation differ from the representation in [3] in that κ is added to the beliefs.

By meta-level reasoning, we have $\mathrm{Con}(\neg\mathsf{B}_b p)$. Since, by assumption, κ does not imply $\mathsf{B}_b p$, we furthermore have $\mathrm{Con}(\kappa \wedge \neg\mathsf{B}_b p)$. The consistency of $\kappa \wedge \neg\mathsf{B}_b p$ is used in an essential way in the derivation (line 6). Let φ denote $\kappa \wedge (\mathsf{b}_a\neg\mathsf{B}_b p \supset \neg\mathsf{B}_b p)$.

1.	$\mathsf{O}_a\varphi \vdash \mathsf{B}_a\varphi$	Def. O
2.	$\mathsf{O}_a\varphi \vdash \mathsf{B}_a\neg\mathsf{B}_a\mathsf{B}_b p \supset \mathsf{B}_a\neg\mathsf{B}_b p$	1, normal logic
3.	$\mathsf{O}_a\varphi \vdash \neg\mathsf{B}_a\mathsf{B}_b p \supset \mathsf{B}_a\neg\mathsf{B}_b p$	2, Æ_I
4.	$\mathsf{O}_a\varphi \vdash \mathsf{C}_a\neg\varphi$	Def. O
5.	$\mathsf{O}_a\varphi \vdash \mathsf{C}_a(\kappa \supset \mathsf{B}_b p)$	4, PL, normal logic
6.	$\mathsf{O}_a\varphi \vdash \Diamond_a(\kappa \wedge \neg\mathsf{B}_b p)$	$\Diamond$-axiom
7.	$\mathsf{O}_a\varphi \vdash \mathsf{C}_a(\kappa \supset \mathsf{B}_b p) \supset \neg\mathsf{B}_a(\kappa \supset \mathsf{B}_b p)$	6, Def $\Diamond$, PL
8.	$\mathsf{O}_a\varphi \vdash \neg\mathsf{B}_a(\kappa \supset \mathsf{B}_b p)$	5, 7, MP
9.	$\mathsf{O}_a\varphi \vdash \neg\mathsf{B}_a\mathsf{B}_b p$	8, normal logic
10.	$\mathsf{O}_a\varphi \vdash \mathsf{B}_a\neg\mathsf{B}_b p$	3, 9, MP

In the third line, we made use of the modal reductive strength of the logic. Notice that a's belief that b does not believe his secret is retracted in case he learns that b has found out that p. I.e. if κ is strengthened such that $\kappa \vdash \mathsf{B}_b p$, line 6 does no longer hold. On the other hand we can prove

$$\mathsf{O}_a(\kappa \wedge (\mathsf{b}_a\neg\mathsf{B}_b p \supset \neg\mathsf{B}_b p)) \vdash \mathsf{B}_a\mathsf{B}_b p.$$

The belief representation in this example is of the same form as the representation in Ex. 1. By the semantic argumentation in Ex. 1, however, we prove the stronger results that $\vdash \mathsf{O}_a\varphi \equiv \mathsf{O}_a(\kappa \wedge \neg\mathsf{B}_b p)$ when $\mathrm{Con}(\kappa \wedge \neg\mathsf{B}_a p)$, and $\vdash \mathsf{O}_a\varphi \equiv \mathsf{O}_a\kappa$ when $\kappa \vdash \mathsf{B}_b p$.

Example 3. In the second example of Halpern and Lakemeyer, they show how one agent reasons about another agent's defeasible reasoning. The letter p stands for "Tweety flies". The scenario is that agent b believes that all agent a knows is that, by default, Tweety flies. It is then derivable that b believes that a believes that Tweety flies. Formally, it is shown that

$$\mathsf{B}_b\mathsf{O}_a(\mathsf{b}_a p \supset p) \vdash \mathsf{B}_b\mathsf{B}_a p.$$

In the same pattern as in the previous example, with $\mathsf{B}_b p$ replaced with $\neg p$ and the $\Diamond$-axiom applied to p, we may show that $\mathsf{O}_a(\mathsf{b}_a p \supset p) \vdash \mathsf{B}_a p$. The desired result is then obtained by normal logic. Notice again that by the reasoning of Ex. 1, we may easily prove the stronger result that $\vdash \mathsf{O}_a(\mathsf{b}_a p \supset p) \equiv \mathsf{O}_a p$.

Example 4. We will in this example continue the the semantic reasoning carried out in Ex. 1. With more than one default there is a chance of mutual conflicts. Let us address the belief representation $\mathsf{O}_a\varphi_2$, where

$$\varphi_2 = \kappa \wedge (\mathsf{b}_a\psi \supset \psi) \wedge (\mathsf{b}_a\gamma \supset \gamma).$$

The content of this belief representation depends on the two default conditionals and the relationship between κ, one the one hand, and ψ and γ on the other hand. The four cases to be analyzed are listed below along with their corresponding equivalents.

1. $\kappa \vdash \neg\psi \wedge \neg\gamma$ $\qquad\vdash \mathsf{O}_a\varphi_2 \equiv \mathsf{O}_a\kappa$
2. $\kappa \vdash \neg\psi$ and $\text{Con}(\kappa, \gamma)$ $\qquad\vdash \mathsf{O}_a\varphi_2 \equiv \mathsf{O}_a(\kappa \wedge \gamma)$
3. $\text{Con}(\kappa, \psi)$, $\text{Con}(\kappa, \gamma)$ and $\kappa \vdash \neg(\psi \wedge \gamma)$ $\qquad\vdash \mathsf{O}_a\varphi_2 \equiv \mathsf{O}_a(\kappa \wedge \psi) \vee \mathsf{O}_a(\kappa \wedge \gamma)$
4. $\text{Con}(\kappa, \psi, \gamma)$ $\qquad\vdash \mathsf{O}_a\varphi_2 \equiv \mathsf{O}_a(\kappa \wedge \psi \wedge \gamma)$

The case where $\kappa \vDash \neg\gamma$ and $\text{Con}(\kappa, \psi)$ is clearly symmetric to the second case and omitted. In the rest of the section, C denotes an arbitrary a-cluster in a saturated model M.

Case 1. By soundness, $\kappa \vDash \neg\psi \wedge \neg\gamma$. Assume first that $M \vDash_C \mathsf{O}_a\varphi_2$. Then $C^+ \subseteq ||\kappa|| \subseteq ||\neg\psi \wedge \neg\gamma||$, and hence (i) $M \vDash_C \mathsf{B}_a\neg\psi \wedge \mathsf{B}_a\neg\gamma$. Let $x \in C$ be any point such that $M \vDash_x \varphi_2$. By PL, $M \vDash_x \kappa$. Conversely, let $x \in C$ be any point such that $M \vDash_x \kappa$. By (i), $M \vDash_x \kappa \wedge \mathsf{B}_a\neg\psi \wedge \mathsf{B}_a\neg\gamma$, and hence $M \vDash_x \varphi_2$ by PL. For the other direction, assume that $M \vDash_C \mathsf{O}_a\kappa$. Then $C^+ = ||\kappa|| \subseteq ||\neg\psi \wedge \neg\gamma||$. Then statement (i) holds, and the proof is as above.

Case 2. Assume first that $M \vDash_C \mathsf{O}_a\varphi_2$. Observe that if $\kappa \vdash \neg\psi$, then $\vdash \varphi_2 \equiv \kappa \wedge \mathsf{B}_a\neg\psi \wedge (\mathsf{b}_a\gamma \supset \gamma)$. Then, since $M \vDash_C \mathsf{O}_a\varphi_2$ and $\kappa \vDash \neg\psi$, $C^+ \subseteq ||\kappa|| \subseteq ||\neg\psi||$. Hence (ii) $M \vDash_C \mathsf{B}_a\neg\psi$. Furthermore, $C^- \subseteq ||\neg\kappa \vee \mathsf{b}_a\psi \vee (\mathsf{b}_a\gamma \wedge \neg\gamma)||$. By (ii), $C^- \subseteq ||\neg\kappa \vee (\mathsf{b}_a\gamma \wedge \neg\gamma)||$, and so (iii) $C^- \subseteq ||\kappa \supset \neg\gamma||$. Since $\text{Con}(\kappa \wedge \gamma)$, $\kappa \wedge \gamma$ is satisfiable. By Lemma 10, there is a point $x \in C$ such that $M \vDash_x \kappa \wedge \gamma$. By (iii), $x \in C^+$, and hence (iv) $M \vDash_C \mathsf{b}_a\gamma$.

Let $x \in C$ be any point such that $M \vDash_x \varphi_2$. By (iv), $M \vDash_x \varphi_2 \wedge \mathsf{b}_a\gamma$. Since $\varphi_2 \vdash \kappa \wedge (\mathsf{b}_a\gamma \supset \gamma)$, $M \vDash_x \kappa \wedge \gamma$. Conversely, assume that $M \vDash_x \kappa \wedge \gamma$. By PL, $M \vDash_x \kappa \wedge (\mathsf{b}_a\gamma \supset \gamma)$. By (ii), $M \vDash_x \kappa \wedge \mathsf{B}_a\neg\psi \wedge (\mathsf{b}_a\gamma \supset \gamma)$, and so $M \vDash_x \varphi_2$.

For the other direction, assume that $M \vDash_C \mathsf{O}_a(\kappa \wedge \gamma)$. Then $C^+ = ||\kappa \wedge \gamma|| \subseteq ||\neg\psi||$, and so (v) $M \vDash_C \mathsf{B}_a\gamma$ and (vi) $M \vDash_C \mathsf{B}_a\neg\psi$. Let $x \in C$ be any point such that $M \vDash_x \varphi_2$. By PL, $M \vDash_x (\kappa \wedge \mathsf{B}_a\neg\gamma) \vee (\kappa \wedge \gamma)$. By (v) it must be the case that $M \vDash_x \kappa \wedge \gamma$. Conversely, let $M \vDash_x \kappa \wedge \gamma$. By (vi), $M \vDash_x \kappa \wedge \mathsf{B}_a\neg\psi \wedge \gamma$. By PL, $M \vDash_x \varphi_2$.

Case 3. Assume first that $M \vDash_C \mathsf{O}_a\varphi_2$. Observe that if $\kappa \vdash \neg(\psi \wedge \gamma)$, then

$$\vdash \varphi_2 \equiv (\kappa \wedge \mathsf{B}_a\neg\psi \wedge (\mathsf{b}_a\gamma \supset \gamma)) \vee (\kappa \wedge \mathsf{B}_a\neg\gamma \wedge (\mathsf{b}_a\psi \supset \psi)).$$

Hence (vii) $M \vDash_C \mathsf{B}_a\neg\psi$ or $M \vDash_C \mathsf{B}_a\neg\gamma$. Furthermore, (viii) $C^- \subseteq ||(\kappa \wedge \psi) \supset \mathsf{b}_a\gamma||$ and $C^- \subseteq ||(\kappa \wedge \gamma) \supset \mathsf{b}_a\psi||$. Since $\text{Con}(\kappa, \psi)$ and $\text{Con}(\kappa, \gamma)$, both $\kappa \wedge \psi$ and $\kappa \wedge \gamma$ are satisfiable. By Lemma 10, there is a point $x \in C$ such that $M \vDash_x \kappa \wedge \psi$, and there is a point $y \in C$ such that $M \vDash_y \kappa \wedge \gamma$. If both formulae are satisfied in C^-, (viii) entails that $M \vDash_C \mathsf{b}_a\psi$ and $M \vDash_C \mathsf{b}_a\gamma$, contradicting (vii). Hence, there is a point $x \in C^+$ such that $M \vDash_x \kappa \wedge \psi$ or there is a point $y \in C^+$ such that $M \vDash_y \kappa \wedge \gamma$. Since $\kappa \vDash \neg(\psi \wedge \gamma)$, there are two cases. As they are symmetrical, we will only treat the first.

If there is a point $x \in C^+$ such that $M \vDash_x \kappa \wedge \psi$, then $M \vDash_x \kappa \wedge \psi \wedge \neg\gamma$. Hence (iv) $M \vDash_C \mathsf{b}_a\psi$, and by (vii), we have (x) $M \vDash_C \mathsf{B}_a\neg\gamma$. Assume first that $M \vDash_x \varphi_2$ for any $x \in C$. By (iv), $M \vDash_x \varphi_2 \wedge \mathsf{b}_a\psi$. By PL, we get $M \vDash_x \kappa \wedge \psi$. Conversely, assume that $M \vDash_x \kappa \wedge \psi$. By (x) $M \vDash_x \kappa \wedge \psi \wedge \mathsf{B}_a\neg\gamma$. $M \vDash_x \varphi_2$ then follows by PL. Similarly, $M \vDash_C \varphi_2 \equiv (\kappa \wedge \gamma)$ if C^+ satisfies $\kappa \wedge \gamma$.

For the other direction, assume that $M \vDash_C \mathsf{O}_a(\kappa \wedge \psi) \vee \mathsf{O}_a(\kappa \wedge \gamma)$. Since $\neg\text{Con}(\kappa, \psi, \gamma)$, $M \vDash_C \mathsf{O}_a(\kappa \wedge \psi)$ implies that $M \vDash_C \neg\mathsf{O}_a(\kappa \wedge \gamma)$. There are hence two cases. We will only treat $M \vDash_C \mathsf{O}_a(\kappa \wedge \psi)$ as the other case is symmetrical. Since $M \vDash_C \mathsf{O}_a(\kappa \wedge \psi)$ and $\kappa \wedge \psi \vDash \neg\gamma$, $C^+ = ||\kappa \wedge \psi|| \subseteq ||\neg\gamma||$.

Hence (xi) $M \vDash_C \mathsf{B}_a \neg\gamma$. Furthermore, (xii) $C^- = ||\kappa \supset \neg\psi||$. Since Con$(\kappa, \psi)$, Lemma 10 entails that there is a point $x \in C$ such that $M \vDash_x \kappa \wedge \psi$. By (xii), $\kappa \wedge \psi$ is satisfied in C^+, and hence (xiii) $M \vDash_C \mathsf{b}_a\psi$. Let $x \in C$ be any point such that $M \vDash_x \varphi_2$. By (xiii), $M \vDash_x \varphi_2 \wedge \mathsf{b}_a\psi$. $M \vDash_x \kappa \wedge \psi$ then follows by PL. Conversely, let $M \vDash_x \kappa \wedge \psi$ for any $x \in C$. By (xi), $M \vDash_x \kappa \wedge \psi \wedge \mathsf{B}_a\neg\gamma$. $M \vDash_x \varphi_2$ then follows by PL.

Case 4. Assume first that $M \vDash_C \mathsf{O}_a\varphi_2$. Then (xiv) $C^- \subseteq ||\kappa \supset \neg(\psi \wedge \gamma)||$. Since Con$(\kappa, \psi, \gamma)$, there is by Lemma 10 a point $x \in C$ such that $M \vDash_x \kappa \wedge \psi \wedge \gamma$. By (xiv) $\kappa \wedge \psi \wedge \gamma$ must be satisfied in C^+, and hence (xv) $M \vDash_C \mathsf{b}_a\psi \wedge \mathsf{b}_a\gamma$. Let $x \in C$ be any point such that $M \vDash_x \varphi_2$. By (xv), $M \vDash_x \varphi_2 \wedge \mathsf{b}_a\psi \wedge \mathsf{b}_a\gamma$. Then $M \vDash_x \kappa \wedge \psi \wedge \gamma$ by PL. If $M \vDash_x \kappa \wedge \psi \wedge \gamma$, $M \vDash_x \varphi_2$ follows immediately by PL. For the other direction, assume that $M \vDash_C \mathsf{O}_a(\kappa \wedge \psi \wedge \gamma)$. Then $C^- = ||\kappa \supset \neg(\psi \wedge \gamma)||$. Argue as above.

5 Finite Languages

In this section we address finite languages, i.e. languages in which the set of propositional letters and the index set I (i.e. the number of different modalities) are finite. We are, moreover, particularly interested in languages bounded by modal depth. To this end let the set $\mathcal{L}_k$ be defined as the set of all formulae φ for which $\mathrm{m}(\varphi) \leq k$. The language of propositional logic is thus denoted $\mathcal{L}_0$ and the set of formula free of modality a with modal depth no more than k is denoted $\mathcal{L}_k^{\setminus a}$.

The language layers are conceptually significant since a belief $\mathsf{B}_a\varphi$ with $\mathrm{m}(\varphi) = k$ is a representation of a representation that agent a has in $\mathcal{L}_k^{\setminus a}$. Moreover, a finite set Γ of formulae has always an upper bound N on modal depth, and an analysis of Γ can be made by studying $\mathcal{L}_N$. This section contains results of this sort.

5.1 Finite Model Property

Technically, Theorem 7 shows that in order to prove that L_I has the Finite model property it is sufficient that we for each φ identify a filtration set Σ such that (i) the filtration $M^\dagger$ of the canonical model is indeed an $\mathrm{m}(\varphi)$-saturated model and (ii) that the size of Σ is bounded by the size of $\mathrm{m}(\varphi)$.

To this end we shall define a finite set of formulae Σ_N such that for every consistent formula φ such that $\mathrm{m}(\varphi) \leq N$, there is a Boolean combination γ of formulae from Σ_N such that $\vdash \varphi \equiv \gamma$.

Assume that the language is finite; let $\{p_1, \ldots, p_n\}$ be the propositional letters. An *atom* is defined as a conjunction $\pm p_1 \wedge \cdots \pm p_n$, where $\pm p_i$ means either p_i or $\neg p_i$. A *molecule* is a disjunction of atoms. Let Γ be a set of formulae. An *a-atom over Γ* is a conjunction of the elements of a maximal consistent subset of $\Gamma^{\setminus a}$, an *a-molecule over Γ* a disjunction of a-atoms over Γ.

By induction we define sets of formulae Σ_k. Σ_0 is defined as the set of atoms. The set Σ_{k+1}^a is defined as the set of formulae that contains $\mathsf{B}_a\mu$ and $\mathsf{C}_a\mu$ for each a-molecule μ over Σ_k, and $\mathsf{b}_a\alpha$ and $\mathsf{c}_a\alpha$ for each a-atom α over Σ_k. Σ_{k+1} is defined as $\Sigma_k \cup \bigcup_{a \in I} \Sigma_{k+1}^a$.

Note that the modal depth of an a-atom over Σ_k is k and that an a-atom over Σ_0 is an atom. Also note that in the inductive step of the definition we only prefix formulae free of modality a in Σ_k with an a-modality. A consequence of this is that every formula in Σ_k is first-order.

Let $\alpha = \bigwedge \Delta$ be an a-atom over Σ_k; Δ a maximal consistent subset of $\Sigma_k^{\setminus a}$. Let $\varphi = \bigwedge \Gamma$ be a consistent formula such that $\Gamma \subseteq \Sigma_k^{\setminus a}$. If $\Gamma \subseteq \Delta$, we will say that α *extends* φ. Note that if $\alpha_1, \ldots, \alpha_m$ represent every extension of φ, then $\vdash \varphi \equiv \alpha_1 \vee \cdots \vee \alpha_m$.

Lemma 11. *Any consistent formula with modal depth $\leq N$ is equivalent to a Boolean combination of formulae of Σ_N.*

Proof. It is easy to see that a purely Boolean formula is equivalent to a molecule. A molecule is a Boolean combination of atoms, and by construction, each atom is contained in Σ_N. We need to prove that given any formula $\mathsf{B}_a\varphi$ with modal depth $\leq N$, there is an equivalent formula $\mathsf{B}_a\varphi' \in \Sigma_N$.

We may without loss of generality assume that $\mathsf{B}_a\varphi$ is first-order, i.e. that φ is free of modality a, and that φ is in disjunctive normal form. Let $k = \mathrm{m}(\varphi)$. We prove by induction on k that φ is equivalent to an a-molecule over Σ_k.

The base case is when $k = 0$, and the result then follows by the initial observation. Suppose inductively that $k > 0$, and let δ be one of the disjuncts of φ. δ is a conjunction of a purely Boolean formula, ($\top$ if δ is completely modalized) and a completely b-modalized formula for each agent $b \neq a$ such that δ is not free of b. We will first transform δ to an equivalent formula δ' by replacing each subformula with depth $< k$ with equivalent formulae with the same modal depth, and then transform δ' to disjunctive normal form.

The purely Boolean conjunct transforms to an equivalent disjunction of atoms. By construction, each atom is contained in Σ_k. A conjunct of the form $\mathsf{B}_b\psi$ is such that, by the induction hypothesis, ψ is equivalent to a b-molecule ψ' over Σ_{k-1}. Let $\mathsf{B}_b\psi'$ be the transformation of $\mathsf{B}_b\psi$ by the substitution of equivalent formulae. The same transformation applies to formulae $\mathsf{C}_b\psi$. By construction, $\mathsf{B}_b\psi'$ and $\mathsf{C}_b\psi'$ are contained in Σ_k. By the same transformation, a conjunct of the form $\mathsf{b}_a\psi$ may be replaced by the equivalent formula $\mathsf{b}_a\psi'$. Let $\psi' = \alpha_1 \vee \cdots \vee \alpha_m$, where each disjunct is a b-atom over Σ_{k-1}. By modal logic, b_b distributes over disjunctions, so $\mathsf{b}_b\psi'$ is equivalent to $\mathsf{b}_b\alpha_1 \vee \cdots \vee \mathsf{b}_b\alpha_m$. By construction, each disjunct is contained in Σ_k. The same transformation applies to formulae $\mathsf{c}_b\psi$. Let δ' be the formula equivalent to δ after these substitutions of equivalent formulae. δ' is now a Boolean combination of formulae free of a in Σ_k. Let δ'' be δ' on disjunctive normal form. Each disjunct β of δ'' is a conjunction of elements of Σ_k. Let $\beta_1, \ldots, \beta_m$ be the a-atoms over Σ_k extending β. The disjunction $\beta_1 \vee \cdots \vee \beta_m$ is then an a-molecule over Σ_k equivalent to β. Since the disjunction of a-molecules over Σ_k clearly is an a-molecule over Σ_k, this transformation of every conjunct of δ'' yields an a-molecule over Σ_k. The initial formula φ is now transformed to an equivalent formula φ', where φ' is a disjunction of a-molecules over Σ_k. φ is thus equivalent to an a-molecule φ' over Σ_k, and by construction, $\mathsf{B}_a\varphi' \in \Sigma_N$. □

Theorem 8. *L_I has the Finite model property.*

Proof. We use the same filtration construction as for $Æ_I$, but while any subformula closed set counts as a filtration set for $Æ_I$ this is not the case for L_I. Any subformula-closed set is sufficient for establishing that $M^\dagger$ is an $Æ_I$-model. However, we want to show that it is also an N-saturated model, and to this end we use $\mathsf{Sf}(\Sigma_N)$ as filtration set. To prove that $M^\dagger$ is N-saturated, we reason similarly to the proof of Lemma 9, using Lemma 11 to establish N-definability of points in the model. The theorem then follows from Theorem 7. □

Decidability follows as a simple corollary to this result. Given that Halpern and Lakemeyer have already showed that their logic (and hence L_I) is PSPACE complete [3], this result is hardly surprising. There are, however, other interesting results of Theorem 8 which we shall address below. First another key notion.

Let $\mathcal{L}^*$ be any sublanguage. We will say that a formula $\varphi \in \mathcal{L}^*$ is a *complete theory* over $\mathcal{L}^*$ iff for all $\psi \in \mathcal{L}^*$, either $\varphi \vdash \psi$ or $\varphi \vdash \neg\psi$.

Theorem 9. *An a-atom over Σ_k is a complete theory over $\mathcal{L}_k^{\setminus a}$.*

Proof. Let α be any a-atom over Σ_k, and let ψ be any formula in $\mathcal{L}_k^{\setminus a}$. By Lemma 11, ψ is equivalent to an a-molecule μ over Σ_k. Since an a-atom over Σ_k is a conjunction of a maximal consistent subset of $\Sigma_k^{\setminus a}$, two a-atoms over Σ_k are consistent iff they are equivalent. As μ is a disjunction of a-atoms over Σ_k, $\alpha \vdash \mu$ if there is a disjunct of μ which is equivalent to α. If there is no such disjunct, α is inconsistent with every disjunct of μ, and so $\alpha \vdash \neg\mu$. $\square$

For $\mathcal{L}_0$, an atom can be seen as a propositional valuation. An atom is then a complete theory in the sense that it provides a complete characterization of the external world with respect to a given agent. The language $\mathcal{L}_k^{\setminus a}$, $k > 0$, on the other hand can be interpreted as a language representing the external world with respect to agent a, where the external world is characterized by a modal language representing the cognitive state of every agent different from a. A complete theory for $\mathcal{L}_k^{\setminus a}$ is then a complete characterization of the external world with respect to agent a.

5.2 The Modal Reduction Theorem

Let β be any formula. The Modal reduction theorem states that there are formulae $\beta_1, \ldots, \beta_m$ free of modality a, such that

$$\vdash \mathsf{O}_a\beta \equiv \mathsf{O}_a\beta_1 \vee \cdots \vee \mathsf{O}_a\beta_m.$$

Moreover, each formula $\mathsf{O}_a\beta_i$ is defined directly from one of the a-clusters satisfying $\mathsf{O}_a\beta$, and each such a-cluster is represented by a formula $\mathsf{O}_a\beta_i$.

Given a formula $\mathsf{O}_a\beta$, let $\mathrm{m}(\mathsf{O}_a\beta) = N$. The filtration set Σ is defined as Σ_N closed under subformulae. Let M^Σ be the filtration of M^c through Σ. Notice that since, by Lemma 11, any formula $\varphi \in \mathcal{L}_N$ is equivalent to a Boolean combination of formulae in Σ_N, if φ is satisfiable, it is satisfied in M^Σ.

For notational simplicity, we will in the following usually denote points in the filtrated model by the letters x and y rather than by the denotation $|s|$ for equivalence classes modulo $\sim_\Sigma$.

It is straightforward to define points and sets of points in the model from an agent a's point of view. Let $x \in W^\Sigma$ and let $X \subseteq W^\Sigma$. Recall from Section 3.2 that $[x]_a$ denotes the equivalence class of x modulo $\approx_a$, and that if $x \approx_a y$, then x and y agree on all formulae free of modality a. As we will define x by the formulae free of modality a that hold at x, this definition defines every point in $[x]_a$. The formula $[\![x]\!]_a$ is defined by the conjunction $[\![x]\!]_a = \bigwedge\{\varphi \mid M^\Sigma \vDash_x \varphi,\ \varphi \in \Sigma^{\setminus a}\}$, while the set of formulae $[\![X]\!]_a$ is defined by $[\![X]\!]_a = \bigvee\{[\![x]\!]_a \mid x \in X\}$. The relation between the equivalence classes modulo $\approx_a$ and the definition of points in the model is summarized by the following result:

Lemma 12. *$M^\Sigma \vDash_y [\![x]\!]_a$ iff $y \approx_a x$ iff $y \in [x]_a$ iff $\vdash [\![x]\!] \equiv [\![y]\!]$.*

Proof. If $M^\Sigma \vDash_y [\![x]\!]_a$, x and y agree on every formula in $\Sigma^{\setminus a}$. The equivalences then follow by definition. $\square$

Let C be an a-cluster of M^Σ and let C^+ and C^- be the belief part and co-belief part, respectively, of C. The *belief formula* of C is defined as $\mathsf{Bel}_C = \bigwedge \mathsf{b}_a[\![x]\!]_a \wedge \mathsf{B}_a([\![C^+]\!]_a)$, where x ranges over the elements of C^+. The *co-belief formula* of C is defined as $\mathsf{Cobel}_C = \bigwedge \mathsf{c}_a[\![x]\!]_a \wedge \mathsf{C}_a([\![C^-]\!]_a)$, where x ranges over the elements of C^-. We state the following two technical results without proof.

Lemma 13. *Each a-cluster C is uniquely characterized by $\mathsf{Bel}_C \wedge \mathsf{Cobel}_C$.*

Lemma 14. *Let C be a bisected a-cluster. Then* $\vdash \mathsf{Bel}_C \wedge \mathsf{Cobel}_C \equiv \mathsf{O}_a[\![C^+]\!]_a$.

Theorem 10. *Let β be any formula. Then* $\vdash \mathsf{O}_a\beta \equiv \mathsf{O}_a\beta_1 \vee \cdots \vee \mathsf{O}_a\beta_m$, *where $\beta_1 \ldots \beta_m$ are formulae free of modality a.*

Proof. Let ψ be the first-order formula equivalent to $\mathsf{O}_a\beta$ and assume that $\mathrm{m}(\psi) = N$. By Lemma 11, ψ is equivalent to a Boolean combination of some formulae in Σ_N. Assume first that $\psi \in s$, $s \in W^c$. As ψ is a Boolean combination of formulae in Σ_N, and M^Σ is the filtration of M^c through $\Sigma = \mathsf{Sf}(\Sigma_N)$, we may apply the Filtration Theorem 3 to infer that $M^\Sigma \vDash_{|s|} \psi$. Let C be the a-cluster containing $|s|$. Since $\vdash \psi \equiv \mathsf{O}_a\beta$, the Soundness Theorem entails that $M^\Sigma \vDash_{|s|} \mathsf{O}_a\beta$. By Lemma 4, C is bisected. Lemma 14 is then applicable, and we may infer that $M^\Sigma \vDash_{|s|} \mathsf{O}_a[\![C^+]\!]_a$. By the Filtration Theorem, $\mathsf{O}_a[\![C^+]\!]_a \in s$. This proves that $\vdash \mathsf{O}_a\beta \supset \mathsf{O}_a[\![C_1^+]\!]_a \vee \cdots \vee \mathsf{O}_a[\![C_m^+]\!]_a$, where $C_1, \ldots, C_m$ are the clusters satisfying $\mathsf{O}_a\beta$.

For the other direction, let C be an a-cluster in M^Σ such that $M^\Sigma \vDash_C \psi$, and let $\mathsf{O}_a[\![C^+]\!]_a \in s$. Let D be the a-cluster containing $|s|$. By the Filtration Theorem, $M^\Sigma \vDash_{|s|} \mathsf{O}_a[\![C^+]\!]_a$, and then also $M^\Sigma \vDash_D \mathsf{O}_a[\![C^+]\!]_a$. By Lemma 4, D is bisected, and Lemma 14 can be used to infer that $M^\Sigma \vDash_D \mathsf{Bel}_C \wedge \mathsf{Cobel}_C$. By Lemma 13, $C = D$, i.e. $|s| \in C$. By assumption, $M^\Sigma \vDash_C \psi$, and so $M^\Sigma \vDash_{|s|} \psi$. By the Filtration Theorem, $\psi \in s$. We have then proved that $\mathsf{O}_a[\![C^+]\!]_a \vdash \psi$. As ψ is equivalent to $\mathsf{O}_a\beta$, we have $\mathsf{O}_a[\![C^+]\!]_a \vdash \mathsf{O}_a\beta$ for every a-cluster C satisfying $\mathsf{O}_a\beta$, i.e. $\mathsf{O}_a[\![C_1^+]\!]_a \vee \cdots \vee \mathsf{O}_a[\![C_m^+]\!]_a \vdash \mathsf{O}_a\beta$, where $C_1, \ldots, C_m$ are the a-clusters satisfying $\mathsf{O}_a\beta$. ☐

Lemma 15. *Let φ be a formula free of modality a. Then there is a unique a-cluster satisfying $\mathsf{O}_a\varphi$.*

Proof. Let C be an a-cluster of M^Σ satisfying $\mathsf{O}_a\varphi$, i.e. every $x \in C^+$ is such that $M^\Sigma \vDash_x \varphi$ and every $x \in C^-$ is such that $M^\Sigma \vDash_x \neg\varphi$. Furthermore, any cluster satisfying $\mathsf{O}_a\beta$ is by Lemma 4 bisected. Hence, C is the unique cluster satisfying $\mathsf{O}_a\varphi$, since for any other bisected cluster D, either there is a point $x \in D^+$ satisfying $\neg\varphi$ or there is a point $x \in D^-$ satisfying φ. ☐

Let $\mathrm{m}(\mathsf{O}_a\varphi) = N$. We will say that $\mathsf{O}_a\varphi$ is an *explicit belief representation* if for any completely a-modalized formula $\psi \in \mathcal{L}_N$, either $\mathsf{O}_a\varphi \vdash \psi$ or $\mathsf{O}_a\varphi \vdash \neg\psi$. In other terms, an explicit belief representation is a formula that determines the agent's attitude toward any formula.

Theorem 11. *Let φ be a formula free of modality a. Then $\mathsf{O}_a\varphi$ is an explicit belief representation.*

Proof. Let C be any a-cluster satisfying $\mathsf{O}_a\varphi$. By Lemma 15, C is the unique cluster satisfying $\mathsf{O}_a\varphi$. Let ψ be any formulae in $\mathcal{L}_N^{\setminus a}$. Either ψ is satisfied by every point in C^+, or there is a point in C^+ satisfying $\neg\psi$. In the first case, $M^\Sigma \vDash_C \mathsf{B}_a\psi$, and in the latter case, $M^\Sigma \vDash_C \neg\mathsf{B}_a\psi$. The same argument applies to any C_a-modal formula. Since we without loss of generality may assume that any completely a-modalized formula $\chi \in \mathcal{L}_N$ is first-order, the result holds for any Boolean combination of completely a-modalized formula. Thus, for any point $x \in W^\Sigma$, if $M^\Sigma \vDash_x \mathsf{O}_a\varphi$, either $M^\Sigma \vDash_x \chi$ or $M^\Sigma \vDash_x \neg\chi$. The theorem now follows by the Filtration Theorem. ☐

Related to the notion of an explicit belief representation is the notion of an *implicit* belief representation, i.e. formulae of the form $\mathsf{O}_a\varphi$ that allow ambiguity with respect to a-modalized formulae. An implicit belief representation is a formula $\mathsf{O}_a\varphi$ where φ is not free of modality a. By applying the Modal reduction theorem, such formulae are reduced to disjunctions of formulae, each of them an explicit belief representation.

6 Summary

In this paper we have defined a standard Kripke semantics for L_I, a multi-modal *only knowing* system. By applying techniques from standard modal logic we have established completeness and the Finite model property for the system. Also, we have showed soundness; by the peculiar nature of the logic, we first have to show consistency, then completeness and finally soundness. By a careful analysis of the filtrated canonical model we have also showed that the logic satisfies an interesting modal reduction property.

The nature of the logic, and hence also its models, is complex and hard to penetrate. It is possible to define a logical space syntactically in the core system $Æ_I$ and replace the $\Diamond$-axiom of L_I with this formula. An advantage of this is that is gives a better insight into the nature of the models; in particular it makes all the options of beliefs explicit. An analysis of this is the provided in the second author's thesis [9] and is the aim of a follow-up paper.

Although the main results have not been previously published, they are contained in the first author's doctoral thesis [10]. The proof of the Modal reduction theorem applies techniques introduced by Segerberg [8] for the single modality logic in response to [10].

References

1. Halpern J.Y. Reasoning about Only Knowing with Many Agents. In *Proc. of the 11th National Conference on Artificial Intelligence (AAAI-93)*, 1993.
2. Halpern J.Y. A Theory of Knowledge and Ignorance for Many Agents. In *Journal of Logic and Computation* 7:1 (79-108), 1997.
3. Halpern J.Y. and Lakemeyer, G. Multi-Agent Only Knowing. In *Journal of Logic and Computation* 11:1 (40-70), 2001.
4. van der Hoek, W. and Jaspars, J A General Approach to Multi-Agent Minimal Knowledge: With Tools and Samples. *Studia Logica* 72 (61-84), 2002.
5. Lakemeyer, G. All They Know: A Study in Multi-Agent Autoepistemic Reasoning. In *Proc. of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93)* (376-381), 1993.
6. Levesque, H. J. All I Know: Study in Autoepistemic Logic. *Artificial Intelligence* 42 (263-309), 1990.
7. Lian, E., Langholm, T. and Waaler, A. *Only Knowing with Confidence Levels: Reductions and Complexity.* In J. Alferes and J. Leite (eds.), Logics in Artificial Intelligence. Ninth European Conference, JELIA 2004, Lisbon, Portugal, September 2004, Proceedings. Vol. 3225 of Lecture Notes in Artificial Intelligence, Springer-Verlag, 2004.
8. Segerberg, K.: Some Modal Reduction Theorems in Autoepistemic Logic. Uppsala Prints and Preprints in Philosophy (1995)
9. Solhaug, B.: Logical spaces in multi-modal Only knowing logics. MSc thesis, University of Oslo, 2004.
10. Waaler, A.: Logical Studies in Complementary Weak S5. Doctoral thesis, University of Oslo (1994)
11. Waaler, A.: Consistency proofs for systems of multi-agent only knowing. To appear in *Advances in Modal Logic 2004.*