

Logical Spaces in Multi-Agent Only Knowing Systems

Bjørnar Solhaug^{1,2} and Arild Waaler^{3,4}

¹ SINTEF ICT, Norway

² Dep. of Information Science and Media Studies, University of Bergen, Norway

³ Finnmark College, Norway

⁴ Dep. of Informatics, University of Oslo, Norway

Abstract. We present a weak multi-agent system of Only knowing and an analysis of the logical spaces that can be defined in it. The logic complements the approach to generalizing Levesque’s All I Know system made by Halpern and Lakemeyer. A novel feature of our approach is that the logic is defined entirely at the object level with no reference to meta-concepts in the definition of the axiom system. We show that the logic of Halpern and Lakemeyer can be encoded in our system in the form of a particular logical space.

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. A variety of multi-modal only knowing languages have been analyzed in [5]. However, to represent defeasible patterns of reasoning in a multi-agent context, only knowing systems in which the underlying modal belief logic is K45 are particularly interesting. A natural way to design such systems is to generalize the only knowing system of Levesque [7] to the multi-modal case.

This is, however, 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 [2–4, 6] 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 [12] the second author introduced another generalization of the only knowing system of Levesque [7] to the multi-modal case which does not use meta-language operators. The \diamond -axiom is instead generalized to the statement that

$\diamond_a\varphi$ is an axiom for each *consistent a-objective* φ (*a-objective* because any occurrence of an *a*-modal operator is within the scope of a *b*-modal operator, $a \neq b$). 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).

We propose yet another solution based on a constructive explication of the concept of a logical space. Intuitively, the function of a logical space is to explicate every logical possibility of the logic. This is what has been done in the single-agent only knowing system $\mathcal{A}\mathcal{E}$ presented in the paper by Lian et al. [8] and further analyzed in [11]. A logical space in $\mathcal{A}\mathcal{E}$ is a formula λ such that for each purely Boolean φ , either $\lambda \vdash \diamond\varphi$ or $\lambda \vdash \neg\diamond\varphi$. Compared to the only knowing system of Levesque, the system $\mathcal{A}\mathcal{E}$ has an increased expressive power due to the possibility of varying the logical space, where the \diamond -axiom of Levesque corresponds to only one of many possibilities.

This paper provides a generalization of the system $\mathcal{A}\mathcal{E}$ to the multi-agent case, and hence also provides a solution to the problem with the generalized \diamond -axiom. In the single-agent case, the set of possibilities is derived from a set of formulae from the language of propositional logic. The formulae of this particular set are referred to as *atoms*. Where α is an atom describing a logical possibility, the logical space λ is defined such that $\diamond\alpha$ is entailed by λ . In the multi-agent case, we aim at defining a logical space λ_a for each agent a , such that $\diamond_a\varphi$ is entailed by λ_a for each φ representing a logical possibility to agent a .

By providing logical spaces as a solution to the problem with the \diamond -axiom, we need not encode meta-concepts into the language, nor refer to such concepts in the definition of the axiom system.

In order to bring the task of defining a multi-agent logical space to a manageable level, we will address the problem inductively at different levels of complexity, each level corresponding to a sub-language within which the set of possibilities is outlined. The base case is equivalent to the single agent case: Let \mathcal{L}_0 denote the language of propositional logic. The set of possibilities is derived by closing a subset of \mathcal{L}_0 under the \diamond_a -operator for each agent a . The resulting set of formulae is then a subset of the language of the next level, denoted \mathcal{L}_1 . Inductively, the set of possibilities for agent a at level $k + 1$ is derived from a subset of *the a-objective formulae of \mathcal{L}_k* .

The main task of this paper is to construct the sets of formulae that, for each agent and each language level, express each and every logical possibility. We propose this as a replacement of the \diamond -axiom of Levesque. In Sect. 4, we will show how the logic of $\mathcal{A}\mathcal{E}_I$ applies to examples from the paper of Halpern and Lakemeyer [4]. In Sect. 6, we prove the equivalence between the systems $\mathcal{A}\mathcal{E}_I$ and L_I , and hence the equivalence between the system of Halpern and Lakemeyer [4], where a particular logical space is added to the axioms of $\mathcal{A}\mathcal{E}_I$.

In [13] a modal reduction property for L_I is established, 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 latter expressions explicitly characterize the possible cognitive states of the agent, given the initial “only knowing” expression. In Sect. 5 we shall see that the same property holds also in \mathcal{AE}_I .

2 The Logic \mathcal{AE}_I

2.1 Syntax

The object language \mathcal{L} contains a countable set of propositional letters \mathcal{P} , the propositional constant \perp , the Boolean connectives \neg and \wedge and the modal operators \mathbf{B}_a and \mathbf{C}_a for each a in a countable non-empty set of indices I . The index set I represents the set of agents, \mathbf{B}_a is a belief operator, and \mathbf{C}_a is a complementary *co-belief* operator for agent $a \in I$. The propositional constant \top is defined as $\neg\perp$, while the Boolean connectives \vee , \supset and \equiv are the usual abbreviations. Other modal operators defined as abbreviations are the following: $\mathbf{b}_a\varphi$ (φ is compatible with belief) is $\neg\mathbf{B}_a\neg\varphi$, $\mathbf{c}_a\varphi$ (φ is compatible with co-belief) is $\neg\mathbf{C}_a\neg\varphi$, $\square_a\varphi$ (φ is necessary) is $\mathbf{B}_a\varphi \wedge \mathbf{C}_a\varphi$ and $\diamond_a\varphi$ (φ is possible) is $\mathbf{b}_a\varphi \vee \mathbf{c}_a\varphi$. Observe that necessity and possibility are relative to the extension of a given agent’s belief and co-belief; the notion of necessity hence captures *personal* necessity.

The more accurate interpretation of the \mathbf{B}_a -operator is that a formula $\mathbf{B}_a\varphi$ states that agent a believes at least φ to be true, but perhaps more. The \mathbf{B}_a -operator thus puts a *lower* bound on the extension of belief. The complementary operator \mathbf{C}_a puts an *upper* bound on the belief in the sense that a formula $\mathbf{C}_a\varphi$ states that agent a believes at most φ to be false, but perhaps less. The formula $\mathbf{B}_a\varphi \wedge \mathbf{C}_a\neg\varphi$ states that φ is *exactly* what is believed. The introduction of the \mathbf{C}_a -operator thus allows an “All a knows”-proposition $\mathbf{O}_a\varphi$ to be defined as $\mathbf{B}_a\varphi \wedge \mathbf{C}_a\neg\varphi$.

A formula not mentioning any modal operators is called *purely Boolean*. φ is an *a-modal atom* if it is of the form $\mathbf{B}_a\varphi$ or $\mathbf{C}_a\varphi$, $a \in I$. An *a-modal literal* is an *a-modal atom* or the negation of an *a-modal atom*. φ is a *completely a-modalized formula* if it is a Boolean combination of *a-modal atoms*. 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 $\mathbf{B}_a\psi$ and $\mathbf{C}_a\psi$ in φ , ψ is free of modality a . If Γ is a set of formulae, $\Gamma^a = \{\varphi \in \Gamma \mid \varphi \text{ free of modality } a\}$ and $\Gamma^a = \{\varphi \in \Gamma \mid \varphi \text{ completely } a\text{-modalized}\}$. If Γ is a set of formulae, $\mathbf{Sf}(\Gamma)$ denotes the set of subformulae of the formulae in Γ . When Γ is a singleton set containing φ , $\mathbf{Sf}(\varphi)$ denotes $\mathbf{Sf}(\{\varphi\})$.

The *modal depth* $\mathbf{d}(\varphi)$ of a formula φ expresses the nesting of alternating modalities in φ . Formally, the modal depth of a purely Boolean φ is 0. Otherwise, if φ is $\mathbf{B}_a\psi$ or $\mathbf{C}_a\psi$, let Ψ be the set of modal atoms which occur as subformulae in ψ . Then $\mathbf{d}(\varphi)$ is the maximal number in $\{\mathbf{d}(\chi) + 1 \mid \chi \in \Psi \text{ and } \chi \text{ is not } a\text{-modalized}\} \cup \{\mathbf{d}(\chi) \mid \chi \in \Psi \text{ and } \chi \text{ is } a\text{-modalized}\}$. Otherwise, the modal depth of φ is the maximal $\mathbf{d}(\psi)$ for a subformula ψ of φ . The

modal depth of an a -modal formula φ is hence increased by prefixing φ with any other modal operator than an a -modal operator.

If Γ is a set of formulae, $\Gamma_k = \{\varphi \in \Gamma \mid \mathbf{d}(\varphi) \leq k\}$. We will in this paper be interested in sub-languages relative to a given modal depth and a given agent. Since \mathcal{L} denotes the language of \mathbb{A}_I (which is just a set of formulae) these sublanguages are denoted \mathcal{L}_k , $\mathcal{L}_k^{\setminus a}$ and \mathcal{L}_k^a following the set indexing notation introduced above.

A *tautology* is a substitution instance of a formula valid in propositional logic, e.g. $\Box_a \varphi \supset \Box_a \varphi$. The deducibility relation ' \vdash ' of the logic \mathbb{A}_I is defined as the least relation that contains all tautologies, is closed under all instances of the rules

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

and contains all instances of the following schemata for each $a \in I$:

$$\begin{array}{ll} K_{\mathbf{B}}: \mathbf{B}_a(\varphi \supset \psi) \supset (\mathbf{B}_a \varphi \supset \mathbf{B}_a \psi) & \overline{B}_{\Box}: \neg \mathbf{B}_a \varphi \supset \Box_a \neg \mathbf{B}_a \varphi \\ K_{\mathbf{C}}: \mathbf{C}_a(\varphi \supset \psi) \supset (\mathbf{C}_a \varphi \supset \mathbf{C}_a \psi) & \overline{C}_{\Box}: \neg \mathbf{C}_a \varphi \supset \Box_a \neg \mathbf{C}_a \varphi \\ B_{\Box}: \mathbf{B}_a \varphi \supset \Box_a \mathbf{B}_a \varphi & T: \Box_a \varphi \supset \varphi \\ C_{\Box}: \mathbf{C}_a \varphi \supset \Box_a \mathbf{C}_a \varphi & \end{array}$$

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

$$\frac{\varphi_1, \dots, \varphi_n \vdash \psi}{\mathbf{B}_a \varphi_1, \dots, \mathbf{B}_a \varphi_n \vdash \mathbf{B}_a \psi}$$

and the corresponding rule for \mathbf{C}_a .

Lemma 1. \Box_a is an S5 modality.

Lemma 2. Any formula is provably equivalent to a first-order formula with the same modal depth.

The former of these two results is Lemma 1 of [13]; the latter is Lemma 2 of [12]. For proofs and further details about the results in the rest of this section, the reader may consult [13].

2.2 Semantics

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 3. E_a is an equivalence relation.

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 4. $C = C^+ \cup C^-$.

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

$$\begin{aligned} \mathcal{M} \models_x p &\leftrightarrow x \in V(p), \quad p \text{ a propositional letter,} \\ \mathcal{M} \models_x \neg\varphi &\leftrightarrow \mathcal{M} \not\models_x \varphi, \\ \mathcal{M} \models_x \mathbf{B}_a\varphi &\leftrightarrow \forall y (xR_ay \rightarrow \mathcal{M} \models_y \varphi), \\ \mathcal{M} \models_x \mathbf{C}_a\varphi &\leftrightarrow \forall y (xS_ay \rightarrow \mathcal{M} \models_y \varphi), \end{aligned}$$

and in the usual way for the other Boolean connectives. We write $\mathcal{M} \models_X \varphi$ iff $(\forall x \in X)(\mathcal{M} \models_x \varphi)$. A formula is valid in a frame if it is true at all points in all models on the frame. If φ is valid in all frames, we write $\models \varphi$, and say that φ is *valid*. $\Gamma \models \varphi$ means that for all models, φ is true at all points which satisfy all formulae in Γ . Note that if C is an *a-cluster*, all points in C agree on every completely *a-modalized* formula in every model on the frame.

Theorem 1. \mathcal{AE}_I is sound, complete and decidable.

Proof. This can be proved by the use of standard techniques from modal logic, see [13]. \square

3 Finitely Bound Sublanguages

The syntax and semantics of the previous section generalize the syntax and semantics of the system \mathcal{AE} to the multi-agent case. What remains to be done is to generalize the notion of a logical space to the multi-agent case. The properties of the single-agent logical space will serve as guiding principles for our multi-agent generalization.

In the single-agent case, under the assumption that there are finitely many propositional letters in the language, say p_1, \dots, p_m , an *atom* is defined as a conjunction $\pm p_1 \wedge \dots \wedge \pm p_m$, where $\pm p$ means either p or $\neg p$. An atom can be interpreted as characterizing the material content of a state of affairs. There are 2^m atoms. Where $\alpha_1, \dots, \alpha_n$ characterize the conceivable states of affairs, the logical space is defined as the formula

$$\diamond \alpha_1 \wedge \dots \wedge \diamond \alpha_n \wedge \Box(\alpha_1 \vee \dots \vee \alpha_n).$$

In the *maximal* logical space, all atoms are possible, i.e. $n = 2^m$.

The notion of a logical space is a notion of personal necessity. To see this, observe that we may define a logical space such that concepts that intuitively are logically independent are related in the logic at the level of necessity. We may e.g. define a logical space λ such that $\lambda \vdash \Box(\text{penguin}(\text{Tweety}) \supset \text{bird}(\text{Tweety}))$.

In order to define a multi-agent logical space, we need to generalize the notion of an atom. To do this, we first define a *finite multi-modal language*. For such a language to be finite, the set of propositional letters and the set of different modalities, i.e. the index set I , obviously need to be finite. Furthermore, as we may construct new formulae by prefixing a formula free of modality a with any a -modal operator, the finite language must be bound by a limit on modal depth. Under the assumption that the set of propositional letters and the set of agents are finite, we will for each modal depth k operate with the sublanguage \mathcal{L}_k .

The notion of an atom is a twin to the notion of a *complete theory*. Given any language \mathcal{L}^* , a formula $\varphi \in \mathcal{L}^*$ is a complete theory for \mathcal{L}^* iff for all formulae $\psi \in \mathcal{L}^*$, either $\varphi \vdash \psi$ or $\varphi \vdash \neg\psi$. The modal language of \mathbb{A} , denoted \mathcal{L}_1 for a singleton set I , can be seen as representations of formulae from \mathcal{L}_0 , i.e. the language of propositional logic. As an atom can be interpreted as a propositional valuation, it is easy to see that each atom is a complete theory for \mathcal{L}_0 . Generalizing this to the multi-agent case is to say that a multi-modal language is a language representing representations in a modal language, the first capturing the cognitive state of a given agent a , the latter representing the material content of states of affairs as well as the cognitive state of every agent different from a . The latter language is then denoted by \mathcal{L}_k^a for some given integer k , the former the closure of \mathcal{L}_k^a under the a -modal operators. A complete theory for \mathcal{L}_k^a is thus a complete characterization of the material content of a state of affairs in addition to a complete characterization of the cognitive state of every agent different from a given agent a .

The notion of a complete theory may then serve as a test for deciding whether a suggested multi-modal logical space is a correct generalization of the single-agent logical space. In other terms, if $\diamond_a \varphi_1 \wedge \dots \wedge \diamond_a \varphi_n \wedge \Box_a(\varphi_1 \vee \dots \vee \varphi_n)$ is a multi-modal logical space for a given agent a , where $\{\varphi_1, \dots, \varphi_n\} \subseteq \mathcal{L}_k^a$ for a given integer k , then each $\varphi \in \{\varphi_1, \dots, \varphi_n\}$ should be a complete theory for \mathcal{L}_k^a .

We will use the following notation for the distribution of a modality over a set of formulae: $\mathbf{B}_a \Gamma = \{\mathbf{B}_a \gamma \mid \gamma \in \Gamma\}$, and the same for any other modality.

Definition 1. Let $\Phi \subseteq \mathcal{L}^a$. The functions Bel_a , Cobel_a and Lspace_a , all of them from a set of formulae free of modality a to a completely a -modalized formula, are defined as follows:

$$\begin{aligned}\text{Bel}_a(\Phi) &= \bigwedge \mathbf{b}_a\Phi \wedge \mathbf{B}_a(\bigvee \Phi), \\ \text{Cobel}_a(\Phi) &= \bigwedge \mathbf{c}_a\Phi \wedge \mathbf{C}_a(\bigvee \Phi), \\ \text{Lspace}_a(\Phi) &= \bigwedge \diamond_a\Phi \wedge \square_a(\bigvee \Phi).\end{aligned}$$

$\text{Lspace}_a(\Phi)$ is the logical space for agent a spanned by Φ . If $\Phi^+ \cup \Phi^- = \Phi$, then $\text{Bel}_a(\Phi^+) \wedge \text{Cobel}_a(\Phi^-)$ is a doxastic a -alternative spanned by Φ . Notice that a doxastic a -alternative spanned by Φ entails $\text{Lspace}_a(\Phi)$. The set of all doxastic a -alternatives spanned by every nonempty subset of Φ is denoted $\text{Dox}_a(\Phi)$.

Two properties will play a central role in our analysis. A set of formulae Φ satisfies the *strong independence property* if every two elements of Φ are consistent iff they are equivalent. Φ is \mathcal{L}^* -saturated if $\Phi \subseteq \mathcal{L}^*$ and every formula $\varphi \in \mathcal{L}^*$ is equivalent to a disjunction of formulae in Φ (\mathcal{L}^* any language addressed in this paper).

Lemma 5. If $\Phi \subseteq \mathcal{L}^a$ satisfies strong independence, then so does $\text{Dox}_a(\Phi)$.

Proof. Let δ_1 and δ_2 be two distinct elements of $\text{Dox}_a(\Phi)$. Then δ_1 and δ_2 must disagree on the belief set or the co-belief set. We treat the former. Let $\delta_1 \vdash \text{Bel}_a(\Gamma_1)$ and $\delta_2 \vdash \text{Bel}_a(\Gamma_2)$. There is then a formula φ such that either $\varphi \in \Gamma_1$ and $\varphi \notin \Gamma_2$ or vice versa. In the first case, $\varphi \wedge \gamma \vdash \perp$ for each $\gamma \in \Gamma_2$ by strong independence. Hence $\varphi \wedge \bigvee \Gamma_2 \vdash \perp$. By modal logic, $\mathbf{B}_a(\bigvee \Gamma_2) \vdash \mathbf{B}_a\neg\varphi$. Since $\delta_2 \vdash \mathbf{B}_a(\bigvee \Gamma_2)$ and $\delta_1 \vdash \mathbf{b}_a\varphi$, we get $\delta_1 \wedge \delta_2 \vdash \perp$. The latter case is symmetrical. \square

Lemma 6. Let \mathcal{L}^* be any Boolean closed set of formulae and Φ be \mathcal{L}^* -saturated. Then $\vdash \bigvee \Phi$.

Proof. Assume that $\neg \bigvee \Phi$ is consistent. Since Φ is \mathcal{L}^* -saturated and \mathcal{L}^* is Boolean closed, there must then be a non-empty set $\Gamma \subseteq \Phi$ such that $\vdash \bigvee \Gamma \equiv \neg \bigvee \Phi$. But this is clearly impossible. \square

Lemma 7. Let Φ be \mathcal{L}_k^a -saturated. Then $\text{Dox}_a(\Phi)$ is \mathcal{L}_{k+1}^a -saturated.

Proof. Let $\varphi \in \mathcal{L}_{k+1}^a$. We may without loss of generality assume that φ is first-order. Since Φ is \mathcal{L}_k^a -saturated, the formulae inside the scope of the a -modalities are equivalent to disjunctions of formulae from Φ . By standard propositional reasoning and normal modal logic and, φ is equivalent to a formula on DNF, where each disjunct is of the form $\psi = \bigwedge \mathbf{b}_a\Gamma_1 \wedge \mathbf{B}_a(\bigvee \Gamma_2) \wedge \bigwedge \mathbf{c}_a\Gamma_3 \wedge \mathbf{C}_a(\bigvee \Gamma_4)$, $\Gamma_1, \dots, \Gamma_4$ subsets of Φ . Let

$$\Delta = \{\delta \in \text{Dox}_a(\Phi) \mid \delta = \text{Bel}_a(\Phi^+) \wedge \text{Cobel}_a(\Phi^-), \Gamma_1 \subseteq \Phi^+ \subseteq \Gamma_2, \Gamma_3 \subseteq \Phi^- \subseteq \Gamma_4\}.$$

Then $\vdash \psi \equiv \bigvee \Delta$. To see that $\bigvee \Delta \vdash \psi$, observe that if $\Gamma_1 \subseteq \Phi^+$, then $\text{Bel}_a(\Phi^+) \vdash \mathbf{b}_a \gamma$ for each $\gamma \in \Gamma_1$, and if $\Phi^+ \subseteq \Gamma_2$, then $\mathbf{B}_a(\bigvee \Phi^+) \vdash \mathbf{B}_a(\bigvee \Gamma_2)$. Conversely, assume that $\psi \not\vdash \bigvee \Delta$, i.e. that ψ is consistent with $\neg \bigvee \Delta$. This entails that ψ is consistent with a formula θ constructed as a conjunction out of the negation of one conjunct from each δ in Δ . But by construction of ψ there is no such θ which is consistent with ψ . \square

Lemma 8. *Let Φ be \mathcal{L}_k^a -saturated and satisfy strong independence, and let δ be a doxastic a -alternative spanned by $\Gamma \subseteq \Phi$. Then δ is a complete theory over \mathcal{L}_{k+1}^a .*

Proof. We need to prove that either $\delta \vdash \psi$ or $\delta \vdash \neg \psi$ for every $\psi \in \mathcal{L}_{k+1}^a$. By Lemma 2, we may without loss of generality assume that ψ is first-order. The result for Boolean combinations of formulae follows easily once the result is established for modal atoms. It suffices to deal with the case where ψ is of the form $\mathbf{B}_a \varphi$, as the other cases are symmetrical.

Let $\delta \vdash \text{Bel}_a(\Gamma_1)$ and $\vdash \varphi \equiv \bigvee \Gamma_2$, Γ_1 and Γ_2 subsets of Φ . There are two cases. Either $\Gamma_1 \subseteq \Gamma_2$, or there is a formula γ such that $\gamma \in \Gamma_1$ and $\gamma \notin \Gamma_2$. In the first case, $\bigvee \Gamma_1 \vdash \bigvee \Gamma_2$. By modal logic, $\mathbf{B}_a(\bigvee \Gamma_1) \vdash \mathbf{B}_a(\bigvee \Gamma_2)$, and so $\delta \vdash \mathbf{B}_a \varphi$. In the second case, $\gamma \wedge \bigvee \Gamma_2 \vdash \perp$ by strong independence. By modal logic, $\mathbf{b}_a \gamma \vdash \mathbf{b}_a \neg(\bigvee \Gamma_2)$. Since $\delta \vdash \mathbf{b}_a \gamma$, we get that $\delta \vdash \mathbf{b}_a \neg(\bigvee \Gamma_2)$, i.e. $\delta \vdash \neg \mathbf{B}_a \varphi$. \square

We are now ready to generalize the single-agent notion of an atom to the multi-agent case. In the single agent case, an atom α can be interpreted as a complete characterization of the material content of a state of affairs. In the multi-agent case, we want for each agent $a_i \in I$, $I = \{a_1, \dots, a_m\}$, and each modal depth k to define a doxastic alternative δ_i , such that δ_i completely characterizes the cognitive state of agent a_i . A conjunction $\alpha \wedge \delta_1 \wedge \dots \wedge \delta_m$ is then a complete characterization of the material content of a state of affairs, as well as a complete characterization of the cognitive state of every agent. As we shall see, the conjunction $\alpha \wedge \delta_1 \wedge \dots \wedge \delta_m$ is a complete theory for \mathcal{L}_k . This conjunction will be referred to as *an I -atom with depth k* .

Given a set of I -atoms with depth k , the doxastic alternatives for agent a with depth $k + 1$ will be defined over this set. Intuitively, where Φ is the set of I -atoms with depth k , the set of formulae $\text{Dox}_a(\Phi)$ is the set of doxastic a -alternatives with depth $k + 1$. This is, however, not the correct generalization of the single agent case, since in the single-agent case, a doxastic alternative is defined over a set of *purely Boolean formulae*. Generalizing this is to define a doxastic alternative for agent a_i over a set of formulae *free of modality a* . To this end, we will define a set of formulae from \mathcal{L}_k^a each formula of which forms a complete theory for \mathcal{L}_k^a .

Convention. Let $\varphi = \alpha \wedge \delta_1 \wedge \dots \wedge \delta_m$ be a formula such that $\alpha \in \mathcal{L}_0$ and δ_i is a doxastic a -alternative. Then $\varphi[a_i/\top] = \alpha \wedge \delta_1 \wedge \dots \wedge \delta_{i-1} \wedge \top \wedge \delta_{i+1} \wedge \dots \wedge \delta_m$. If Φ is a set of I -atoms, $\Phi[a_i/\top] = \{\varphi[a_i/\top] \mid \varphi \in \Phi\}$.

Definition 2 (I-atoms). The set of I-atoms Φ_k with depth k is defined as follows: Φ_0 is the set of atoms, while Φ_{k+1} is all formulae $\alpha \wedge \delta_1 \wedge \dots \wedge \delta_m$ such that

- α is an atom,
- δ_i is a doxastic a_i -alternative spanned by the set $\Gamma_i \subseteq \Phi_k[a_i/\top]$,
- $\exists \varphi \in \Phi_k$ such that $\varphi \vdash \alpha$ and for each a_i , $\varphi[a_i/\top] \in \Gamma_i$.

From now on Φ_k refers to the set of I-atoms with depth k . The third condition in the definition above is a consistency condition as witnessed by the following result.

Lemma 9. Assume that each Φ_k satisfies strong independence and that $\Phi_k[a/\top]$ is \mathcal{L}_k^a -saturated for each agent a . Let α be an atom and δ_i be a doxastic a_i -alternative spanned by $\Gamma_i \subseteq \Phi_k[a_i/\top]$. Then $\alpha \wedge \delta_1 \wedge \dots \wedge \delta_m$ is consistent if and only if $\exists \varphi \in \Phi$ such that $\varphi \vdash \alpha$ and for each a_i , $\varphi[a_i/\top] \in \Gamma_i$.

Proof. Note that if δ_i is spanned by the set $\Gamma_i \subseteq \Phi_k[a_i/\top]$, then $\delta_i \vdash \Box_{a_i}(\bigvee \Gamma_i)$. By axiom T , $\delta_i \vdash \bigvee \Gamma_i$. Since the conjuncts of ψ are of different modalities (the atom purely Boolean, however), inconsistency of ψ can stem from axiom T only. Hence, it suffices to prove that the consistency condition ensures consistency of $\psi = \alpha \wedge \bigvee \Gamma_1 \wedge \dots \wedge \bigvee \Gamma_m$.

Also note that Φ_0 is the set of atoms (which is trivially \mathcal{L}_0 -saturated) and that the condition for $k = 1$ then simply states that there is an atom α such that $\alpha \in \Gamma_i$ for each Γ_i . If there is a Γ_i such that $\alpha \notin \Gamma_i$, Lemma 6 gives that $\bigvee \Gamma_i \vdash \neg\alpha$. Hence ψ is inconsistent. Conversely, if ψ is inconsistent, there must be a Γ_i such that $\alpha \notin \Gamma_i$, and hence the condition is not satisfied.

If $k > 1$, suppose that the condition is not satisfied. Then, for each $\varphi \in \Phi_k$ which entails α there is a Γ_i such that $\varphi[a_i/\top] \notin \Gamma_i$. It follows from this that given any two distinct sets Γ_i and Γ_j , each two elements $\varphi_1[a_i/\top] \in \Gamma_i$ and $\varphi_2[a_j/\top] \in \Gamma_j$ must disagree on a doxastic a_l -alternative, $a_i \neq a_j \neq a_l$. In other terms, there are two distinct doxastic a_l -alternatives δ_1 and δ_2 such that $\varphi_1[a_i/\top] \vdash \delta_1$ and $\varphi_2[a_j/\top] \vdash \delta_2$. By the strong independence assumption, $\delta_1 \wedge \delta_2 \vdash \perp$, and so $\varphi_1[a_i/\top] \wedge \varphi_2[a_j/\top] \vdash \perp$. Since this holds for any two distinct sets Γ_i and Γ_j , ψ must be inconsistent.

Suppose conversely that ψ is inconsistent. There are two cases. In the first case, there is a set Γ_i such that $\bigvee \Gamma_i \vdash \neg\alpha$, i.e. for each $\varphi \in \Phi_k$ such that $\varphi \vdash \alpha$, there is a set Γ_i such that $\varphi[a_i/\top] \notin \Gamma_i$. Then the condition is not satisfied. In the second case, there are two distinct sets Γ_i and Γ_j such that $\bigvee \Gamma_i \wedge \bigvee \Gamma_j \vdash \perp$. Then, for each two elements $\varphi_1[a_i/\top] \in \Gamma_i$ and $\varphi_2[a_j/\top] \in \Gamma_j$, $\varphi_1[a_i/\top] \wedge \varphi_2[a_j/\top] \vdash \perp$. We may assume that α is entailed by both φ_1 and φ_2 since this was treated in the first case. Since $\varphi_1[a_i/\top] \wedge \varphi_2[a_j/\top] \vdash \perp$, the two formulae must disagree on a doxastic a_l -alternative, and hence φ_1 and φ_2 are two distinct elements of Φ_k . The condition is then not satisfied. \square

Lemma 10. *The set $\Phi_k[a/\top]$ satisfies strong independence.*

Proof. The base case is when $k = 0$. $\Phi_0[a/\top]$ is the set of atoms, and it is immediate that the set of atoms satisfies strong independence. Suppose inductively that φ and ψ are two distinct elements of $\Phi_{k+1}[a/\top]$. Then φ and ψ either disagree on an atom or on a doxastic b -alternative, $b \neq a$. In the first case, it is immediate that $\varphi \wedge \psi \vdash \perp$. In the second case, let $\varphi \vdash \delta_b^1$ and $\psi \vdash \delta_b^2$, where δ_b^1 and δ_b^2 are doxastic b -alternatives spanned by Γ_1 and Γ_2 , respectively, Γ_1 and Γ_2 subsets of $\Phi_k[b/\top]$. By the induction hypothesis, $\Phi_k[b/\top]$ satisfies strong independence. By Lemma 5, $\delta_b^1 \wedge \delta_b^2 \vdash \perp$. Hence $\varphi \wedge \psi \vdash \perp$. \square

Corollary 1. *The set of doxastic a -alternatives spanned by subsets of $\Phi_k[a/\top]$ satisfies strong independence.*

Proof. Immediate from Lemma 10 and Lemma 5. \square

Lemma 11. *$\Phi_k[a/\top]$ is $\mathcal{L}_k^{\setminus a}$ -saturated and $\text{Dox}_a(\Phi_k[a/\top])$ is \mathcal{L}_{k+1}^a -saturated.*

Proof. Both properties are proved by simultaneous induction on k . In the base case $\Phi_0 = \Phi_0^a$. It is easy to see that the first condition holds. Since Φ_0 is \mathcal{L}_0 -saturated, the second holds by Lemma 7.

$\Phi_{k+1}[a/\top]$ is $\mathcal{L}_{k+1}^{\setminus a}$ -saturated (induction step). We have to prove that for each $\varphi \in \mathcal{L}_{k+1}^{\setminus a}$ there is a subset of $\Phi_{k+1}[a/\top]$ the disjunction of which is equivalent to φ . It is easy to see (using the DNF equivalent of each formula) that it is sufficient to prove this for φ of the form $\varphi^{\mathcal{P}} \wedge \varphi^{a_1} \wedge \cdots \wedge \varphi^{a_m}$ where $\varphi^{\mathcal{P}}$ is purely Boolean, φ^a is \top and every other φ^{a_i} is in $\mathcal{L}_{k+1}^{a_i}$. Let the atom set $\widehat{\varphi}^{\mathcal{P}}$ be the set of atoms which imply $\varphi^{\mathcal{P}}$, $\widehat{\varphi}^a$ be $\{\top\}$ and $\widehat{\varphi}^{a_i}$ be the set of all $\delta \in \text{Dox}_{a_i}(\Phi_k[a_i/\top])$ such that $\delta \vdash \varphi$. Let $\widehat{\varphi}$ be the set of every formula $\alpha \wedge \delta_1 \wedge \cdots \wedge \delta_m$ in $\Phi_{k+1}[a/\top]$ such that $\alpha \in \widehat{\varphi}^{\mathcal{P}}$ and $\delta_i \in \widehat{\varphi}^{a_i}$.

It follows by construction that $\bigvee \widehat{\varphi} \vdash \varphi$. Conversely, assume that φ is consistent with $\neg \bigvee \widehat{\varphi}$. By induction hypothesis and Lemma 6, $\vdash \text{Dox}_{a_i}(\Phi_k[a_i/\top])$. This entails that there must be a consistent ψ of the form $\alpha \wedge \delta_1 \wedge \cdots \wedge \delta_m$ which implies φ and which is not in $\Phi_{k+1}[a/\top]$. But this is only possible if δ violates the third subcondition in the definition of Φ_{k+1} (Definition 2). By Lemma 9, ψ is inconsistent. Contradiction. Hence $\vdash \varphi \equiv \bigvee \widehat{\varphi}$.

$\text{Dox}_a(\Phi_k[a/\top])$ is \mathcal{L}_{k+1}^a -saturated (induction step). By the induction hypothesis, $\Phi_k[a/\top]$ is $\mathcal{L}_k^{\setminus a}$ -saturated. Then, by Lemma 7, $\text{Dox}_a(\Phi_k[a/\top])$ is \mathcal{L}_{k+1}^a -saturated. \square

Theorem 2. *Each formula $\varphi \in \Phi_k^{\setminus a}$ is a complete theory over $\mathcal{L}_k^{\setminus a}$. Each doxastic alternative $\delta \in \text{Dox}_a(\Phi_k[a/\top])$ is a complete theory over \mathcal{L}_{k+1}^a .*

Proof. By Lemma 10, $\Phi_k[a/\top]$ satisfies strong independence and by Lemma 11 $\Phi_k[a/\top]$ is $\mathcal{L}_k^{\setminus a}$ -saturated. By Lemma 8, each $\delta \in \text{Dox}_a(\Phi_k[a/\top])$ is a complete theory over \mathcal{L}_{k+1}^a . Since each $\delta \in \text{Dox}_a(\Phi_{k-1}[a/\top])$ is a complete theory over \mathcal{L}_k , it follows that each $\varphi \in \Phi_k[a/\top]$ is a complete theory over $\mathcal{L}_k^{\setminus a}$. \square

Having defined the set of I -atoms, we may now define the logical space for the multi-agent case. A *logical space of agent a up to depth k* is defined over a subset Γ of $\Phi_k[a/\top]$ by the formula $\text{Lspace}_a(\Gamma)$. Observe that for $k = 0$, the logical space is defined by the formula $\text{Lspace}_a(\Delta)$, $\Delta \subseteq \Phi_0$, which is a logical space as defined for the single-agent system \mathcal{A} .

Corollary 2. *Let λ be a logical space for agent a up to k and $\varphi \in \mathcal{L}_k^{\lambda a}$. Then either $\lambda \vdash \diamond_a \varphi$ or $\lambda \vdash \neg \diamond_a \varphi$.*

4 Examples

In Section 6.1 of [4], Halpern and Lakemeyer give examples of how their logic can be used to represent default reasoning in a multi-agent situation. We will show how the inferences are carried out in the logic \mathcal{A}_I .

Example 1. The first example of Halpern and Lakemeyer is this. Let p be agent a 's secret and suppose he makes the assumption that unless he believes that b knows his secret, he assumes that she does not know it. We will now prove that if this is all he believes and if it is conceivable that b does not know his secret, then he believes that she does not know his secret. Formally, we show

$$\lambda_a \wedge \mathbf{O}_a(\neg \mathbf{B}_a \mathbf{B}_b p \supset \neg \mathbf{B}_b p) \vdash \mathbf{B}_a \neg \mathbf{B}_b p,$$

where λ_a is the logical space of agent a . Let φ denote $\neg \mathbf{B}_a \mathbf{B}_b p \supset \neg \mathbf{B}_b p$. Note that the assumption that it is conceivable to a that b does not know his secret implies that $\lambda_a \vdash \diamond_a \neg \mathbf{B}_b p$. Let us turn to the formal derivation.

- | | |
|--|-------------------------------|
| 1. $\lambda_a \wedge \mathbf{O}_a \varphi \vdash \mathbf{B}_a \varphi$ | PL |
| 2. $\lambda_a \wedge \mathbf{O}_a \varphi \vdash \mathbf{C}_a \neg \varphi$ | PL |
| 3. $\lambda_a \wedge \mathbf{O}_a \varphi \vdash (\mathbf{B}_a \varphi \wedge \neg \mathbf{B}_a \mathbf{B}_b p) \supset \mathbf{B}_a \neg \mathbf{B}_b p$ | normal logic, \mathcal{A}_I |
| 4. $\lambda_a \wedge \mathbf{O}_a \varphi \vdash \mathbf{C}_a \neg \varphi \supset (\mathbf{C}_a \neg \mathbf{B}_a \mathbf{B}_b p \wedge \mathbf{C}_a \mathbf{B}_b p)$ | normal logic |
| 5. $\lambda_a \wedge \mathbf{O}_a \varphi \vdash \diamond_a \neg \mathbf{B}_b p$ | assumption |
| 6. $\lambda_a \wedge \mathbf{O}_a \varphi \vdash \mathbf{C}_a \mathbf{B}_b p \supset \neg \mathbf{B}_a \mathbf{B}_b p$ | 5, PL |
| 7. $\lambda_a \wedge \mathbf{O}_a \varphi \vdash \neg \mathbf{B}_a \mathbf{B}_b p$ | 2, 4, 6, PL |
| 8. $\lambda_a \wedge \mathbf{O}_a \varphi \vdash \mathbf{B}_a \neg \mathbf{B}_b p$ | 1, 3, 7, PL |

In the third line, we made use of the modal reductive strength of the logic. The critical point in the derivation is of course the fifth line. This theorem rests on the assumption that $\neg \mathbf{B}_b p$ is conceivable to agent a . The derivation in the system of Halpern and Lakemeyer is somewhat longer, since they need to apply some extra machinery to reason about validity and satisfiability.

The nonmonotonicity becomes apparent when we add $\mathbf{B}_b p$ to the belief set of agent a , or we define the logical space such that $\lambda_a \vdash \neg \diamond_a \neg \mathbf{B}_b p$. Then $\mathbf{B}_a \mathbf{B}_b p$ is deducible.

Example 2. In their next example, Halpern and Lakemeyer show how one agent reasons about another agent's ability to reason nonmonotonically. The letter p

stands for “Tweety flies”. It is then shown that if a believes that all b believes is that by default Tweety flies, then a believes that b believes that Tweety flies.

Again, it is the logical space that makes the deduction go through in our system. But here, since a reasons about b ’s ability to reason nonmonotonically, if a is to derive the conclusion that b believes p , a must believe that the conceivability space of b is such that p is conceivable. I.e., the logical space λ_b of b must be such that $\lambda_b \vdash \diamond_b p$. Note that since the nonmonotonicity in this example is about b , we need not consult the logical space of a .

What we want to prove is thus that

$$\mathbf{B}_a(\lambda_b \wedge \mathbf{O}_b(\mathbf{b}_b p \supset p)) \vdash \mathbf{B}_a \mathbf{B}_b p.$$

In the same pattern as in the previous example, we may show that $\lambda_b \wedge \mathbf{O}_b(\mathbf{b}_b p \supset p) \vdash \mathbf{B}_b p$, given the assumption that $\lambda_b \vdash \diamond_b p$. The difference is that we now reason about agent b and that $\neg \mathbf{B}_b p$ is replaced with p . The desired result then follows by normal logic.

The assumption we made that a believes λ_b is a stronger assumption than what we actually need. It may very well be such that a believes that b has one of several different conceivability spaces. The assumption we need is that every logical space of agent b compatible with a ’s beliefs must be such that $\diamond_b p$ is implied by it.

5 The Modal Reduction Theorem

We will in this section assume a sub-language \mathcal{L}_k bound by a finite set of propositional letters \mathcal{P} , a finite set of indices I and a given modal depth k . Let the logical space λ of agent a be given, and let β be any formula. The modal reduction theorem states that there are formulae β_1, \dots, β_n free of modality a , such that

$$\lambda \vdash \mathbf{O}_a \beta \equiv \mathbf{O}_a \beta_1 \vee \dots \vee \mathbf{O}_a \beta_n.$$

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

Let $\lambda \wedge \mathbf{O}_a \varphi$ have depth k . We will say that $\lambda \wedge \mathbf{O}_a \varphi$ is an *explicit belief representation* if for any formula $\psi \in \mathcal{L}_k^a$, either $\lambda \wedge \mathbf{O}_a \varphi \vdash \psi$ or $\lambda \wedge \mathbf{O}_a \varphi \vdash \neg \psi$. In other terms, an explicit belief representation is a formula that determines the agent’s attitude towards any formula in the language.

Theorem 3. *Let φ be any formula free of modality a . Then $\lambda \wedge \mathbf{O}_a \varphi$ is an explicit belief representation.*

Related to the notion of an explicit belief representation is the notion of an *implicit belief representation*, i.e. formulae of the form $\lambda \wedge \mathbf{O}_a \varphi$ that allow ambiguity with respect to a -modalized formulae. An implicit belief representation is a formula $\lambda \wedge \mathbf{O}_a \varphi$ 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 Related Work

We will in this section prove the equivalence of the system \mathcal{AE}_I with two earlier attempts of generalizing the system of Levesque [7]. The first of these other systems is the system \mathcal{HL} of Halpern and Lakemeyer [4], where a generalization of Levesque's system is provided by coding the satisfiability relation into the system. Notice that the language of \mathcal{HL} is an extension of \mathcal{L} . We will prove the equivalence with \mathcal{AE}_I with respect to the common part of the languages. The second system is the system L_I of Waaler [12], where the \diamond -axiom of Levesque's system is generalized to the statement that $\diamond_a\varphi$ is a theorem provided that φ is a consistent formula free of modality a .

The deducibility relations of \mathcal{HL} and L_I are denoted $\vdash_{\mathcal{HL}}$ and \vdash_{L_I} , respectively. In [12], the equivalence of L_I and \mathcal{HL} was established. We will in this section prove the equivalence of the system \mathcal{AE}_I and the system L_I . The equivalences of the three systems then follow as a corollary.

6.1 The system L_I

Let \vdash' be the deducibility relation given by removing the axiom schema T from the system \mathcal{AE}_I . The deducibility relation \vdash_{L_I} of the system L_I is defined as the least relation extending \vdash' containing every instance of the following schema for each agent $a \in I$:

$$\diamond_a : \diamond_a\varphi \text{ provided } \varphi \not\vdash_{L_I} \perp, \varphi \text{ free of modality } a.$$

There is a circular pattern to the \diamond_a -axiom, but in [12], it is shown that the circularity is not vicious. This result is captured by Lemma 12 below.

As in \mathcal{AE}_I any formula is provably equivalent to a first-order formula in L_I . Moreover, T is a theorem of L_I . Hence, L_I is an extension of \mathcal{AE}_I . For proof of these claims consult [13].

6.2 Equivalence of \mathcal{AE}_I and L_I

L_I is a proper extension of \mathcal{AE}_I . However, equivalence between the systems can be established for sublanguages up to a given depth by strengthening \mathcal{AE}_I with a particular set of formulae. In the single-agent case, when the maximal logical space is added to the axioms of \mathcal{AE} , the system \mathcal{AE} is equivalent to the propositional fragment of Levesque's system. What we need to do in the multi-agent case is to identify a set of formulae that, when added to the axioms of \mathcal{AE}_I , yields equivalence of \mathcal{AE}_I and L_I .

Definition 3 (Maximal I -atoms). *The set of maximal I -atoms with depth k is defined as follows: Φ_0 is the set of atoms, while Φ_{k+1} is all formulae $\alpha \wedge \delta_1 \wedge \dots \wedge \delta_n$ such that*

- α is an atom,
- δ_i is a doxastic a_i -alternative spanned by $\Phi_k[a_i/\top]$,

The critical difference between the definition of a maximal I -atom and the definition of an I -atom as defined in Definition 2 is that δ_i in the inductive step of the definition of a maximal I -atom is spanned by $\Phi_k[a_i/\top]$, and *not subsets* of $\Phi_k[a_i/\top]$. The consistency condition is furthermore omitted. This is because formulae $\alpha \wedge \delta_1 \wedge \dots \wedge \delta_m$ trivially satisfy the consistency condition in the definition of the maximal I -atoms. (We omit the easy proof of this claim.)

The *maximal logical space* of agent a_i with depth k is now defined as $\lambda_i = \text{Lspace}_a(\Phi_{k-1}[a_i/\top])$. We will prove that the set of formulae $\Lambda = \{\lambda_i \mid a_i \in I\}$ added to the axioms of \mathcal{AE}_I yields equivalence with L_I up to depth k .

Before we proceed, we need an important result from [12]. This result states that L_I -consistency of a formula φ free of modality a is established without reference to the theorem $\diamond_a \varphi$.

Lemma 12. *Let φ be L_I -provable. Then there is an L_I -proof π of φ such that $d(\psi) < d(\varphi)$ for every instance of an axiom $\diamond_a \psi$ which is used in π .*

Theorem 4. *Let Λ be the set of maximal logical spaces with depth k for each agent $a_i \in I$ and $d(\varphi) \leq k$. Then $\Lambda \vdash \varphi$ iff $\vdash_{L_I} \varphi$.*

Proof. The proof is by induction on the depth of the logical spaces, and both directions are proved simultaneously. As $\vdash \subseteq \vdash_{L_I}$, we need for the 'only if' direction to prove that $\vdash_{L_I} \bigwedge \Lambda$. For the 'if' direction, we need to prove that L_I is a strengthening of \mathcal{AE}_I by $\bigwedge \Lambda$ only. That is, we need to prove that $\diamond_{a_i} \varphi$ is deducible in \mathcal{AE}_I from Λ , where $\diamond_{a_i} \varphi$ is derivable in L_I by an application of \diamond_{a_i} to a formula φ , where $d(\varphi) < d(\lambda_i)$.

The base case is when each λ_i is spanned by the set of atoms Φ_0 . 'Only if': As every atom α is L_I -consistent, $\vdash_{L_I} \diamond_{a_i} \alpha$ by the \diamond_{a_i} -axiom, and since $\bigvee \Phi_0$ is a PL-tautology, we get $\vdash_{L_I} \Box_{a_i} (\bigvee \Phi_0)$ by RN. So $\vdash_{L_I} \lambda_i$ for every $\lambda_i \in \Lambda$. 'If': Suppose $\vdash_{L_I} \diamond_{a_i} \varphi$ is deduced in L_I by an application of \diamond_{a_i} . It must then be the case that φ is a purely Boolean formula such that $\varphi \not\vdash_{L_I} \perp$. Since L_I extends \mathcal{AE}_I , $\varphi \not\vdash \perp$. There is then an atom α such that $\alpha \vdash \varphi$. By modal logic, $\diamond_{a_i} \alpha \vdash \diamond_{a_i} \varphi$, and so $\lambda_i \vdash \diamond_{a_i} \varphi$.

In the inductive step, let $d(\lambda_i) = k + 1$, λ_i spanned by $\Phi_k[a_i/\top]$. 'Only if': We need to establish that $\psi \not\vdash_{L_I} \perp$ for every $\psi \in \Phi_k[a_i/\top]$ and that $\vdash_{L_I} \bigvee \Phi_k[a_i/\top]$. Once these two properties are established, we may apply \diamond_{a_i} to the first and RN to the latter to get the desired result.

Note that ψ is a conjunction of an atom and a doxastic a_j -alternative δ_j for each $a_j \neq a_i$. Each δ_j entails the maximal logical space λ'_j , $d(\delta_j) = k$. Let Λ' be the set of maximal logical spaces with depth k for each $a_j \neq a_i$. By construction of the logical space, we have $\psi \wedge \Lambda' \not\vdash \perp$. By the induction hypothesis, $\psi \not\vdash_{L_I} \perp$. By axiom \diamond_{a_i} , we get $\vdash_{L_I} \diamond_{a_i} \psi$.

Let Δ_j be the doxastic a_j -alternatives spanned by $\Phi_{k-1}[a_j/\top]$. Observe that for each $\delta_j \in \Delta_j$, $d(\delta_j) = k$ and $\delta_j \vdash \lambda'_j$, where λ'_j is the maximal logical space with depth k for agent a_j . Notice that the set of conjunctions of an atom and a formula $\delta_j \in \Delta_j$ for each $a_j \neq a_i$ is exactly the set of formulae $\Phi_k[a_j/\top]$. In order to prove $\vdash_{L_I} \bigvee \Phi_k[a_i/\top]$, we will prove that $\vdash_{L_I} \bigvee \Delta_j$ for each $a_j \neq a_i$.

The result then follows by standard propositional reasoning and the fact that $\vdash_{L_I} \bigvee \Phi_0$.

We will first prove that $\lambda'_j \vdash \bigvee \Delta_j$. Suppose that $\lambda'_j \not\vdash \bigvee \Delta_j$, i.e. $\lambda'_j \wedge \neg(\bigvee \Delta_j) \not\vdash \perp$. By Lemma 11, there is a doxastic a_j -alternative δ'_j with depth k such that $\delta'_j \vdash \lambda'_j \wedge \neg(\bigvee \Delta_j)$. But then $\delta'_j \vdash \lambda'_j$, and so $\delta'_j \in \Delta_j$. Contradiction. Since $\lambda'_j \vdash \bigvee \Delta_j$, we get $\vdash_{L_I} \bigvee \Delta_j$ by the induction hypothesis. $\vdash_{L_I} \bigvee \Phi_k[a_i/\top]$ follows by standard propositional reasoning, and $\vdash_{L_I} \Box_{a_i}(\bigvee \Phi_k[a_i/\top])$ by RN.

'If': Suppose $\vdash_{L_I} \Diamond_{a_i} \varphi$, $\mathbf{d}(\varphi) < \mathbf{d}(\lambda_i)$, $\lambda_i \in \Lambda$, is deduced in L_I by an application of \Diamond_{a_i} . It must then be the case that φ is a formula free of modality a_i such that $\varphi \not\vdash_{L_I} \perp$. By Lemma 12, any application of the \Diamond_{a_i} -axiom to establish the consistency of φ is to formulae with depth $< \mathbf{d}(\varphi)$. By the induction hypothesis, $\Lambda' \wedge \varphi \not\vdash \perp$, where Λ' is the set of maximal logical spaces with depth k for each $a_j \neq a_i$.

We may without loss of generality assume that φ is first-order and on DNF. Since $\Lambda' \wedge \varphi \not\vdash \perp$, there is a disjunct ψ of φ such that $\Lambda' \wedge \psi \not\vdash \perp$. ψ is a conjunction of a purely Boolean formula ψ^P and a completely a_j -modalized formula ψ^{a_j} for each $a_j \neq a_i$. Since $\lambda'_j \wedge \psi^{a_j} \not\vdash \perp$, $\lambda'_j \in \Lambda'$, there is by Lemma 11 a doxastic a_j -alternative δ_j with depth k such that $\delta_j \vdash \lambda'_j \wedge \psi^{a_j}$. Let Δ be the set of these formulae δ_j for each $a_j \neq a_i$. As to ψ^P , there is an atom α such that $\alpha \vdash \psi^P$. Since each δ_j entails the maximal logical space, the consistency condition is trivially satisfied, and so $\alpha \wedge \Delta \not\vdash \perp$. Since each element of $\{\alpha\} \cup \Delta$ entails a respective conjunct of ψ , we have $\alpha \wedge \Delta \vdash \psi$, and so $\alpha \wedge \Delta \vdash \varphi$. Observe that the conjunction $\alpha \wedge \bigwedge \Delta$ is an element of $\Phi_k[a_j/\top]$ and that $\Diamond_{a_i}(\alpha \wedge \bigwedge \Delta)$ is a conjunct of the maximal logical space λ_i with depth k . Since $\alpha \wedge \Delta \vdash \varphi$, we have $\Diamond_{a_i}(\alpha \wedge \bigwedge \Delta) \vdash \Diamond_{a_i} \varphi$ by modal logic, and so $\lambda_i \vdash \Diamond_{a_i} \varphi$ as desired. \square

Corollary 3. $\Lambda \vdash \varphi$ iff $\vdash_{L_I} \varphi$ iff $\vdash_{HL} \varphi$, $\varphi \in \mathcal{L}$, provided $\mathbf{d}(\varphi) \leq \mathbf{d}(\lambda_i)$ for each $\lambda_i \in \Lambda$.

Proof. Follows immediately from Theorem 15 of [12] and Theorem 4. \square

7 Conclusion and Future Work

The focus of this paper is on the logical foundation of multi-agent systems. We have successfully developed a notion of logical space for agents in a multi-modal only knowing language. Clearly, a practical application will require a more economical way of representing and reasoning within logical spaces, typically achieved by means of highly restricted languages. However, to implement constraints like this, one needs to know what “all the options” are. This paper presents an answer to this fundamental and conceptually important question.

A number of interesting questions can be raised on the basis of this logical clarification. First, we have not presented any complexity analysis. The size of a logical space grows quickly beyond any tractable level. However, in a particular situation one will not need to span the entire space syntactically, exactly like one in \mathcal{AE} can provide an implicit definition of a logical space by means of a characteristic formula [8, 11]. We plan to address this question in a subsequent

paper. We also plan to extend the reduction method used to give a constructive proof of the Modal Reduction Theorem in \mathcal{AE} to \mathcal{AE}_I and to extend the language with language constructs to express different degrees of confidence for each agent (like in \mathcal{AE}). The latter task is in itself straightforward; however, a non-trivial use of this would be to develop a theory of multi-agent default reasoning within this language which generalizes the encoding of default logic in \mathcal{AE} [1].

References

1. Engan, I., Lian, E. H., Langholm, T. and Waaler, A.: Default Reasoning with Preference within Only Knowing Logic. To appear in Proceedings of the 8th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'05), Springer Verlag Lecture Notes in Artificial Intelligence series (2005)
2. Halpern, J. Y.: Reasoning about Only Knowing with Many Agents. Proceedings of the 11th National Conference on Artificial Intelligence (AAAI-93) (1993)
3. Halpern, J. Y.: A Theory of Knowledge and Ignorance for many agents. *Journal of Logic and Computation* **7:1** (1997) 79–108
4. Halpern, J. Y. and Lakemeyer, G.: Multi-Agent Only Knowing. *Journal of Logic and Computation* **11:1** (2001) 40–70
5. Hoek, W. and Thijsse, E.: A General Approach to Multi-Agent Minimal Knowledge: With Tools and Samples. *Studia Logica* **72:1** (2002) 61–84
6. Lakemeyer, G.: All They Know: A Study in Multi-Agent Autoepistemic Reasoning. Proc. of the 13th International Joint Conference on Artificial Intelligence (IJCAI-93) (1993) 376–381
7. Levesque, H. J.: All I know: A study in autoepistemic logic. *Artificial Intelligence* **42** (1990) 263–309
8. Lian, E. H., Angle, T. and Waaler, A.: Only Knowing with Confidence Levels: Reductions and Complexity. JELIA 2004, Proceedings, Lecture Notes in Computer Science **3229** (2004) 500–512
9. Solhaug, B.: Logical Spaces in Multi-Modal Only Knowing Logics. Master's Thesis, University of Oslo (2004)
10. Waaler, A.: Logical Studies in Complementary Weak S5. Doctoral thesis, University of Oslo (1994)
11. Waaler, A., Klüwer, J.W., Langholm, T. and Lian, E.: Only Knowing with Degrees of Confidence. Submitted for publication, 2005.
12. Waaler, A.: Consistency proofs for systems of multi-agent only knowing. To appear in *Advances in Modal Logic* 2005.
13. Waaler, A. and Solhaug, B.: Semantics for Multi-Agent Only Knowing (extended abstract). To appear in the Proceedings of Theoretical Aspects of Rationality and Knowledge (TARK X), 2005.