The Thirty-Second AAAI Conference
on Artificial Intelligence (AAAI-18)

In Praise of Belief Bases: Doing Epistemic Logic without Possible Worlds

Emiliano Lorini
CNRS-IRIT, Toulouse University, France

Abstract

We introduce a new semantics for a logic of explicit and implicit beliefs based on the concept of multi-agent belief base. Differently from existing Kripke-style semantics for epistemic logic in which the notions of possible world and doxastic/epistemic alternative are primitive, in our semantics they are non-primitive but are defined from the concept of belief base. We provide a complete axiomatization and a decidability result for our logic.

Introduction

Epistemic logic and, more generally, formal epistemology are the areas at the intersection between philosophy (Hintikka 1962), artificial intelligence (AI) (Fagin et al. 1995; Meyer and van der Hoek 1995) and economics (Lismont and Mongin 1994) devoted to the formal representation of epistemic attitudes of agents including belief and knowledge. An important distinction in epistemic logic is between *explicit belief* and *implicit belief*. According to (Levesque 1984), "...a sentence is explicitly believed when it is actively held to be true by an agent and implicitly believed when it follows from what is believed" (p. 198). This distinction is particularly relevant for the design of resource-bounded agents who spend time to make inferences and do not believe all facts that are deducible from their actual beliefs.

The concept of explicit belief is tightly connected with the concept of *belief base* (Nebel 1992; Makinson 1985; Hansson 1993; Rott 1998). In particular, an agent's belief base, which is not necessarily closed under deduction, includes all facts that are explicitly believed by the agent. Nonetheless, existing logical formalizations of explicit and implicit beliefs (Levesque 1984; Fagin and Halpern 1987) do not clearly account for this connection.

The aim of this paper is to fill this gap by providing a multi-agent logic that precisely articulates the distinction between explicit belief, as a fact in an agent's belief base, and implicit belief, as a fact that is deducible from the agent's explicit beliefs, given the agents' common ground. The concept of *common ground* (Stalnaker 2002) corresponds to the body of information that the agents commonly believe to be the case and that has has to be in the deductive closure of their belief bases. The multi-agent aspect of the logic lies in the fact that it supports reasoning about agents' high-order beliefs, i.e., an agent's explicit (or implicit) belief about the explicit (or implicit) belief of another agent.

Differently from existing Kripke-style semantics for epistemic logic in which the notions of possible world and doxastic/epistemic alternative are primitive, in the semantics of our logic the notion of doxastic alternative is defined from — and more generally grounded on — the concept of belief base.

We believe that an explicit representation of agents' belief bases is crucial in order to facilitate the task of designing intelligent systems such as robotic agents or conversational agents. The problem of extensional semantics for epistemic logic, whose most representative example is the Kripkean semantics, is their being too abstract and too far from the agent specification. More generally, the main limitation of the Kripkean semantics is that it does not say from where doxastic alternatives come from thereby being ungrounded.[1]

The paper is organized as follows. We first present the language of our logic of explicit and implicit beliefs. Then, we introduce a semantics for this language based on the notion of multi-agent belief base. We also consider two additional Kripke-style semantics in which the notion of doxastic alternative is primitive. These additional semantics will be useful for proving completeness and decidability of our logic. We show that the three semantics are all equivalent with respect to the formal language under consideration. Then, we provide an axiomatization for our logic of explicit and implicit belief and prove that its satisfiability problem is decidable. After having discussed related work, we conclude.

A Language for Explicit and Implicit Beliefs

LDA (Logic of Doxastic Attitudes) is a logic for reasoning about explicit beliefs and implicit beliefs of multiple agents. Assume a countably infinite set of atomic propositions $Atm = \{p, q, \ldots\}$ and a finite set of agents $Agt = \{1, \ldots, n\}$.

[1] The need for a grounded semantics for doxastic/epistemic logics has been pointed out by other authors including (Lomuscio, Qu, and Raimondi 2015).

We define the language of the logic LDA in two steps. We first define the language $\mathcal{L}_0(Atm)$ by the following grammar in Backus-Naur Form (BNF):

$$\alpha \quad ::= \quad p \mid \neg\alpha \mid \alpha_1 \wedge \alpha_2 \mid \mathtt{E}_i\alpha$$

where p ranges over Atm and i ranges over Agt. $\mathcal{L}_0(Atm)$ is the language for representing explicit beliefs of multiple agents. The formula $\mathtt{E}_i\alpha$ is read "agent i explicitly (or actually) believes that α is true". In this language, we can represent high-order explicit beliefs, i.e., an agent's explicit belief about another agent's explicit beliefs.

The language $\mathcal{L}(Atm)$, extends the language $\mathcal{L}_0(Atm)$ by modal operators of implicit belief and is defined by the following grammar:

$$\varphi \quad ::= \quad \alpha \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \mathtt{I}_i\varphi$$

where α ranges over $\mathcal{L}_0(Atm)$. For notational convenience we write $\mathcal{L}_0$ instead of $\mathcal{L}_0(Atm)$ and $\mathcal{L}$ instead of $\mathcal{L}(Atm)$, when the context is unambiguous.

The other Boolean constructions $\top$, $\bot$, $\vee$, $\rightarrow$ and $\leftrightarrow$ are defined from α, $\neg$ and $\wedge$ in the standard way.

For every formula $\varphi \in \mathcal{L}$, we write $Atm(\varphi)$ to denote the set of atomic propositions of type p occurring in φ. Moreover, for every set of formulas $X \subseteq \mathcal{L}$, we define $Atm(X) = \bigcup_{\varphi\in X} Atm(\varphi)$.

The formula $\mathtt{I}_i\varphi$ has to be read "agent i implicitly (or potentially) believes that φ is true". We define the dual operator $\widehat{\mathtt{I}}_i$ as follows:

$$\widehat{\mathtt{I}}_i\varphi \stackrel{\mathtt{def}}{=} \neg\mathtt{I}_i\neg\varphi.$$

$\mathtt{I}_i\varphi$ has to be read "φ is compatible with agent i's implicit beliefs".

Formal Semantics

In this section, we present three formal semantics for the language of explicit and implicit beliefs defined above. In the first semantics, the notion of doxastic alternative is not primitive but it is defined from the primitive concept of belief base. The second semantics is a Kripke-style semantics, based on the concept of notional doxastic model, in which an agent's set of doxastic alternatives coincides with the set of possible worlds in which the agent's explicit beliefs are true. The third semantics is a weaker semantics, based on the concept of *quasi*-notional doxastic model. It only requires that an agent's set of doxastic alternatives has to be included in the set of possible worlds in which the agent's explicit beliefs are true. At a later stage in the paper, we will show that three semantics are equivalent with respect to the formal language under consideration.

Multi-agent belief base semantics

We first consider the semantics based on the concept of multi-agent belief base that is defined as follows.

Definition 1 (Multi-agent belief base) *A multi-agent belief base is a tuple $B = (B_1, \ldots, B_n, V)$ where:*

- *for every $i \in Agt$, $B_i \subseteq \mathcal{L}_0$ is agent i's belief base,*
- *$V \subseteq Atm$ is the actual state.*

A similar concept is used in belief merging (Koieczny and Pérez 2002) in which each agent is identified with her belief base. Our concept of multi-agent belief base also includes the concept of actual state, as the set of true atomics facts.

The sublanguage $\mathcal{L}_0(Atm)$ is interpreted with respect to multi-agent belief bases, as follows.

Definition 2 (Satisfaction relation) *Let $B = (B_1, \ldots, B_n, V)$ be a multi-agent belief base. Then:*

$$\begin{aligned} B \models p &\iff p \in V \\ B \models \neg\alpha &\iff B \not\models \alpha \\ B \models \alpha_1 \wedge \alpha_2 &\iff B \models \alpha_1 \text{ and } B \models \alpha_2 \\ B \models \mathtt{E}_i\alpha &\iff \alpha \in B_i \end{aligned}$$

The following definition introduces the concept of doxastic alternative.

Definition 3 (Doxastic alternatives) *Let $B = (B_1, \ldots, B_n, V)$ and $B' = (B'_1, \ldots, B'_n, V')$ be two multi-agent belief bases. Then, $B\mathcal{R}_iB'$ if and only if, for every $\alpha \in B_i$, $B' \models \alpha$.*

$B\mathcal{R}_iB'$ means that B' is a doxastic alternative for agent i at B (i.e., at B agent i considers B' possible). The idea of the previous definition is that B' is a doxastic alternative for agent i at B if and only if, B' satisfies all facts that agent i explicitly believes at B.

A multi-agent belief model (MAB) is defined to be a multi-agent belief base supplemented with a set of multi-agent belief bases, called *context*. The latter includes all multi-agent belief bases that are compatible with the agents' common ground (Stalnaker 2002), i.e., the body of information that the agents commonly believe to be the case.

Definition 4 (Multi-agent belief model) *A multi-agent belief model (MAB) is a pair (B, Cxt), where B is a multi-agent belief base and Cxt is a set of multi-agent belief bases. The class of MABs is denoted by* $\mathbf{MAB}$.

Note that in the previous definition we do not require $B \in Cxt$. Let us illustrate the concept of MAB with the aid of an example.

Example 1 *Let $Agt = \{1, 2\}$ and $Atm = \{p, q\}$. Moreover, let (B_1, B_2, V) such that:*

$$\begin{aligned} B_1 &= \{p, \mathtt{E}_2p\}, \\ B_2 &= \{p\}, \\ V &= \{p, q\}. \end{aligned}$$

Suppose that the agents have in their common ground the fact $p \rightarrow q$. In other words, they commonly believe that p implies q. This means that:

$$Cxt = \{B' : B' \models p \rightarrow q\}.$$

The following definition generalizes Definition 2 to the full language $\mathcal{L}(Atm)$. Its formulas are interpreted with respect to MABs. (Boolean cases are omitted, as they are defined in the usual way.)

Definition 5 (Satisfaction relation (cont.)) *Let $(B, Cxt) \in \mathbf{MAB}$. Then:*

$$\begin{aligned} (B, Cxt) \models \alpha &\iff B \models \alpha \\ (B, Cxt) \models \mathtt{I}_i\varphi &\iff \forall B' \in Cxt : \text{if } B\mathcal{R}_iB' \text{ then} \\ &\qquad (B', Cxt) \models \varphi \end{aligned}$$

Let us go back to the example.

Example 2 *It is to check that the following holds:*

$$(B, Cxt) \models \mathtt{I}_1(p \wedge q) \wedge \mathtt{I}_2(p \wedge q) \wedge \mathtt{I}_1\mathtt{I}_2(p \wedge q).$$

Indeed, we have:

$$\mathcal{R}_1(B) \cap Cxt = \{B' : B' \models p \wedge \mathtt{E}_2 p \wedge (p \rightarrow q)\},$$
$$\mathcal{R}_2(B) \cap Cxt = \{B' : B' \models p \wedge (p \rightarrow q)\},$$

and, consequently,

$$\big(\mathcal{R}_1 \circ \mathcal{R}_2(B)\big) \cap Cxt = \{B' : B' \models p \wedge (p \rightarrow q)\},$$

where $\circ$ is the composition operation between binary relations and $\mathcal{R}_i(B) = \{B' : B\mathcal{R}_i B'\}$.

Here, we consider consistent MABs that guarantee consistency of the agents' belief bases. Specifically:

Definition 6 (Consistent MAB) *(B, Cxt) is a consistent MAB (CMAB) if and only if, for every $B' \in Cxt \cup \{B\}$, there exists $B'' \in Cxt$ such that $B'\mathcal{R}_i B''$. The class of CMABs is denoted by* $\mathbf{CMAB}$.

Let $\varphi \in \mathcal{L}$, we say that φ is valid for the class of CMABs if and only if, for every $(B, Cxt) \in \mathbf{CMAB}$, we have $(B, Cxt) \models \varphi$. We say that φ is satisfiable for the the class of CMABs if and only if $\neg\varphi$ is not valid for the the class of CMABs.

Notional doxastic model semantics

Let us now consider the semantics for LDA based on the concept of notional doxastic model (NDM). It is defined in the next Definition 7, together with the satisfaction relation for the formulas of the language $\mathcal{L}(Atm)$.

Definition 7 (Doxastic model) *A notional doxastic model (NDM) is a tuple $M = (W, \mathcal{D}, \mathcal{N}, \mathcal{V})$ where:*

- *W is a set of worlds,*
- *$\mathcal{D} : Agt \times W \longrightarrow 2^{\mathcal{L}_0}$ is a doxastic function,*
- *$\mathcal{N} : Agt \times W \longrightarrow 2^W$ is a notional function, and*
- *$\mathcal{V} : Atm \longrightarrow 2^W$ is a valuation function,*

and that satisfies the following conditions for all $i \in Agt$ and $w \in W$:

(C1) $\mathcal{N}(i, w) = \bigcap_{\alpha \in \mathcal{D}(i,w)} ||\alpha||_M$, and

(C2) there exists $v \in W$ such that $v \in \mathcal{N}(i, w)$,

with:

$$\begin{array}{rcl}
(M, w) \models p & \Longleftrightarrow & w \in \mathcal{V}(p) \\
(M, w) \models \neg\varphi & \Longleftrightarrow & (M, w) \not\models \varphi \\
(M, w) \models \varphi \wedge \psi & \Longleftrightarrow & (M, w) \models \varphi \text{ and } (M, w) \models \psi \\
(M, w) \models \mathtt{E}_i\alpha & \Longleftrightarrow & \alpha \in \mathcal{D}(i, w) \\
(M, w) \models \mathtt{I}_i\varphi & \Longleftrightarrow & \forall v \in \mathcal{N}(i, w) : (M, v) \models \varphi
\end{array}$$

and

$$||\alpha||_M = \{v \in W : (M, v) \models \alpha\}.$$

The class of notional doxastic models is denoted by $\mathbf{NDM}$.

We say that a NDM $M = (W, \mathcal{A}, \mathcal{D}, \mathcal{N}, \mathcal{V})$ is *finite* if and only if W, $\mathcal{D}(i, w)$ and $\mathcal{V}^{-1}(w)$ are finite sets for every $i \in Agt$ and for every $w \in W$, where $\mathcal{V}^{-1}$ is the inverse function of $\mathcal{V}$.

For every agent i and world w, $\mathcal{D}(i, w)$ denotes agent i's set of explicit beliefs at w.

The set $\mathcal{N}(i, w)$, used in the interpretation of the implicit belief operator $\mathtt{I}_i$, is called agent i's set of *notional* worlds at world w. The term 'notional' is taken from (Dennett 1996; 1988) (see, also, (Konolige 1986)): an agent's notional world is *a world in which all the agent's explicit beliefs are true*. This idea is clearly expressed by the Condition C1. According to the Condition C2, an agent's set of notional worlds must be non-empty. This guarantees consistency of the agent's implicit beliefs.

Let $\varphi \in \mathcal{L}$, we say that φ is valid for the the class of NDMs if and only if, for every $M = (W, \mathcal{A}, \mathcal{D}, \mathcal{N}, \mathcal{V}) \in \mathbf{NDM}$ and for every $w \in W$, we have $(M, w) \models \varphi$. We say that φ is satisfiable for the the class of NDMs if and only if $\neg\varphi$ is not valid for the the class of NDMs.

Quasi-model semantics

In this section we provide an alternative semantics for the logic LDA based on a more general class of models, called quasi-notional doxastic models (quasi-NDMs). This semantics will be fundamental for proving completeness of LDA.

Definition 8 (Quasi-notional doxastic model) *A quasi-notional doxastic model (quasi-NDM) is a tuple $M = (W, \mathcal{D}, \mathcal{N}, \mathcal{V})$ where $W, \mathcal{D}, \mathcal{N}$ and $\mathcal{V}$ are as in Definition 7 except that Condition C1 is replaced by the following weaker condition, for all $i \in Agt$ and $w \in W$:*

(C1) $\mathcal{N}(i, w) \subseteq \bigcap_{\alpha \in \mathcal{D}(i,w)} ||\alpha||_M$.*

The class of quasi-notional doxastic models is denoted by $\mathbf{QNDM}$. Truth conditions of formulas in $\mathcal{L}$ relative to this class are the same as truth conditions of formulas in $\mathcal{L}$ relative to the class $\mathbf{NDM}$. Validity and satisfiability of a LDA formula φ for the the class of quasi-NDMs are defined in the usual way.

As for NDMs, we say that a quasi-NDM $M = (W, \mathcal{D}, \mathcal{N}, \mathcal{V})$ is *finite* if and only if W, $\mathcal{D}(i, w)$ and $\mathcal{V}^{-1}(w)$ are finite sets for every $i \in Agt$ and for every $w \in W$.

Equivalences between semantics

The present section is devoted to present equivalences between the different semantics for the language $\mathcal{L}(Atm)$. The results of the section are summarized in Figure 1.

The figure highlights that the five semantics for the language $\mathcal{L}(Atm)$ defined in the previous section are all equivalent, as from every node in the graph we can reach all other nodes.

Equivalence between quasi-NDMs and finite quasi-NDMs We use a filtration argument to show that if a formula φ of the language $\mathcal{L}$ is true in a (possibly infinite) quasi-NDM then it is true in a finite quasi-NDM.

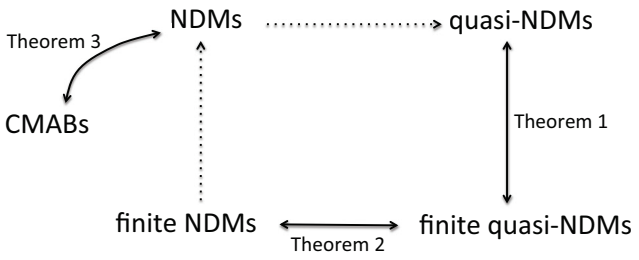

Figure 1: Relations between semantics. An arrow means that satisfiability relative to the first class of structures implies satisfiability relative to the second class of structures. Dotted arrows denote relations that follow straightforwardly given the inclusion between classes of structures.

Let $M = (W, \mathcal{D}, \mathcal{N}, \mathcal{V})$ be a (possibly infinite) quasi-NDM and let $\Sigma \subseteq \mathcal{L}$ be an arbitrary finite set of formulas which is closed under subformulas. (Cf. Definition 2.35 in (Blackburn, de Rijke, and Venema 2001) for a definition of subformulas closed set of formulas.) Let the equivalence relation $\equiv_\Sigma$ on W be defined as follows. For all $w, v \in W$:

$$w \equiv_\Sigma v \text{ iff } \forall \varphi \in \Sigma : (M, w) \models \varphi \text{ iff } (M, v) \models \varphi.$$

Let $|w|_\Sigma$ be the equivalence class of the world w with respect to the equivalence relation $\equiv_\Sigma$.

We define W_Σ to be the filtrated set of worlds with respect to Σ:

$$W_\Sigma = \{|w|_\Sigma : w \in W\}.$$

Clearly, W_Σ is a finite set.

Let us define the filtrated valuation function $\mathcal{V}_\Sigma$. For every $p \in Atm$, we define:

$$\mathcal{V}_\Sigma(p) = \{|w|_\Sigma : (M, w) \models p\} \qquad \text{if } p \in Atm(\Sigma)$$
$$\mathcal{V}_\Sigma(p) = \emptyset \qquad \text{otherwise}$$

The next step in the construction consists in defining the filtrated doxastic function. For every $i \in Agt$ and for every $|w|_\Sigma \in W_\Sigma$, we define:

$$\mathcal{D}_\Sigma(i, |w|_\Sigma) = \mathcal{D}(i, w) \cap \Sigma.$$

Finally, for every $i \in Agt$ and for every $|w|_\Sigma \in W_\Sigma$, we define agent i's set of notional worlds at $|w|_\Sigma$ as follows:

$$\mathcal{N}_\Sigma(i, |w|_\Sigma) = \{|v|_\Sigma : v \in \mathcal{N}(i, w)\}.$$

We call the model $M_\Sigma = (W_\Sigma, \mathcal{A}_\Sigma, \mathcal{D}_\Sigma, \mathcal{N}_\Sigma, \mathcal{V}_\Sigma)$ the filtration of M under Σ.

We can state the following filtration lemma.

Lemma 1 *Let $\varphi \in \Sigma$ and let $w \in W$. Then, $(M, w) \models \varphi$ if and only if $(M_\Sigma, |w|_\Sigma) \models \varphi$.*

PROOF. The proof is by induction on the structure of φ. For the ease of exposition, we prove our result for the language $\mathcal{L}$ in which the "diamond" operator $\widehat{\mathtt{I}}_i$ is taken as primitive and the "box" operator $\mathtt{I}_i$ is defined from it. Since the two operators are inter-definable, this does not affect the validity of our result.

The case $\varphi = p$ is immediate from the definition of $\mathcal{V}_\Sigma$. The boolean cases $\varphi = \neg\psi$ and $\varphi = \psi_1 \wedge \psi_2$ follow straightforwardly from the fact that Σ is closed under subformulas. This allows us to apply the induction hypothesis.

Let us prove the case $\varphi = \mathtt{E}_i\alpha$.

($\Rightarrow$) Suppose $(M, w) \models \mathtt{E}_i\alpha$ with $\mathtt{E}_i\alpha \in \Sigma$. Thus, $\alpha \in \mathcal{D}(i, w)$. Hence, by definition of $\mathcal{D}_\Sigma(i, |w|_\Sigma)$ and the fact that Σ is closed under subformulas, we have $\alpha \in \mathcal{D}_\Sigma(i, |w|_\Sigma)$. It follows that $(M_\Sigma, |w|_\Sigma) \models \mathtt{E}_i\alpha$.

($\Leftarrow$) For the other direction, suppose $(M_\Sigma, |w|_\Sigma) \models \mathtt{E}_i\alpha$ with $\mathtt{E}_i\alpha \in \Sigma$. Thus, $\alpha \in \mathcal{D}_\Sigma(i, |w|_\Sigma)$. Hence, by definition of $\mathcal{D}_\Sigma(i, |w|_\Sigma)$, $\alpha \in \mathcal{D}(i, w)$.

Let us conclude the proof for the case $\varphi = \widehat{\mathtt{I}}_i\psi$. It is easy to check that $\mathcal{N}_\Sigma$ gives rise to the smallest filtration and that the following two properties hold for all $w, v \in W$ and for all $i \in Agt$:

(i) if $v \in \mathcal{N}(i, w)$ then $|v|_\Sigma \in \mathcal{N}_\Sigma(i, |w|_\Sigma)$, and

(ii) if $|v|_\Sigma \in \mathcal{N}_\Sigma(i, |w|_\Sigma)$ then for all $\widehat{\mathtt{I}}_i\varphi \in \Sigma$, if $M, v \models \varphi$ then $M, w \models \widehat{\mathtt{I}}_i\varphi$.

($\Rightarrow$) Suppose $(M, w) \models \widehat{\mathtt{I}}_i\psi$ with $\widehat{\mathtt{I}}_i\psi \in \Sigma$. Thus, there exists $v \in \mathcal{N}(i, w)$ such that $(M, v) \models \psi$. By the previous item (i), $|v|_\Sigma \in \mathcal{N}_\Sigma(i, |w|_\Sigma)$. Since Σ is closed under subformulas, we have $\psi \in \Sigma$. Thus, by the induction hypothesis, $(M_\Sigma, |v|_\Sigma) \models \psi$. It follows that $(M_\Sigma, |w|_\Sigma) \models \widehat{\mathtt{I}}_i\psi$.

($\Leftarrow$) For the other direction, suppose $(M_\Sigma, |w|_\Sigma) \models \widehat{\mathtt{I}}_i\psi$ with $\widehat{\mathtt{I}}_i\psi \in \Sigma$. Thus, there exists $|v|_\Sigma \in \mathcal{N}_\Sigma(i, |w|_\Sigma)$, such that $(M_\Sigma, |v|_\Sigma) \models \psi$. Since Σ is closed under subformulas, by the induction hypothesis, we have $(M, v) \models \psi$. By the item (ii) above, it follows that $(M, w) \models \widehat{\mathtt{I}}_i\psi$. ■

The next step consists in proving that M_Σ is the right model construction.

Proposition 1 *The tuple $M_\Sigma = (W_\Sigma, \mathcal{A}_\Sigma, \mathcal{D}_\Sigma, \mathcal{N}_\Sigma, \mathcal{V}_\Sigma)$ is a finite quasi-NDM.*

PROOF. Clearly, M_Σ is finite. Moreover, it is easy to verify that it satisfies the Condition C2 in Definition 7. We are going to prove that it satisfies the Condition C1* in Definition 8.

By Lemma 1, if $\alpha \in \mathcal{D}(i, w) \cap \Sigma$ then $||\alpha||_{M_\Sigma} = \{|v|_\Sigma : v \in ||\alpha||_M\}$. Moreover, as M is a quasi-NDM, we have

$$\mathcal{N}(i, w) \subseteq \bigcap_{\alpha \in \mathcal{D}(i,w)} ||\alpha||_M \subseteq \bigcap_{\alpha \in \mathcal{D}(i,w) \cap \Sigma} ||\alpha||_M.$$

Hence, by definition of $\mathcal{N}_\Sigma(i, |w|_\Sigma)$ and $\mathcal{D}_\Sigma$,

$$\mathcal{N}_\Sigma(i, |w|_\Sigma) \subseteq \bigcap_{\alpha \in \mathcal{D}_\Sigma(i, |w|_\Sigma)} ||\alpha||_{M_\Sigma}.$$

■

The following is our first result about equivalence between the semantics in terms of quasi-NDMs and the semantics in terms of finite quasi-NDMS.

Theorem 1 *Let $\varphi \in \mathcal{L}$. Then, if φ is satisfiable for the class of quasi-NDMs, if and only if it is satisfiable for the class of finite quasi-NDMs.*

PROOF. The right-to-left direction is obvious. As for the left-to-right direction, let M be a possibly infinite quasi-NDM and let w be a world in M such that $(M, w) \models \varphi$. Moreover, let $sub(\varphi)$ be the set of subformulas of φ. Then, by Lemma 1 and Proposition 1, $(M_{sub(\varphi)}, |w|_{sub(\varphi)}) \models \varphi$ and $M_{sub(\varphi)}$ is a finite quasi-NDM. ■

Equivalence between finite NDMs and finite quasi-NDMs As the following theorem highlights, the LDA semantics in terms of finite NDMs and the LDA semantics in terms of finite quasi-NDMs are equivalent.

Theorem 2 *Let $\varphi \in \mathcal{L}$. Then, φ is satisfiable for the class of finite NDMs if and only if φ is satisfiable for the class of finite quasi-NDMs.*

PROOF. The left-to-right direction is obvious. We are going to prove the right-to-left direction.

Let $M = (W, \mathcal{D}, \mathcal{N}, \mathcal{V})$ be a finite quasi-NDM that satisfies φ, i.e., there exists $w \in W$ such that $(M, w) \models \varphi$. Let

$$\mathcal{T}(M) = \cup_{w \in W, i \in Agt} Atm(\mathcal{D}(i, w))$$

be the *terminology* of model M including all atomic propositions that are in the explicit beliefs of some agent at some world in M. Since M is finite, $\mathcal{T}(M)$ is finite too.

Let us introduce an injective function:

$$f : Agt \times W \longrightarrow Atm \setminus (\mathcal{T}(M) \cup Atm(\varphi))$$

which assigns an identifier to every agent in Agt and world in W. The fact that Atm is infinite while W, $\mathcal{T}(M)$ and $Atm(\varphi)$ are finite guarantees that such an injection exists.

The next step consists in defining the new model $M' = (W', \mathcal{D}', \mathcal{N}', \mathcal{V}')$ with $W' = W$, $\mathcal{N}' = \mathcal{N}$ and where $\mathcal{D}'$ and $\mathcal{V}'$ are defined as follows.

For every $i \in Agt$ and for every $w \in W$:

$$\mathcal{D}'(i, w) = \mathcal{D}(i, w) \cup \{f(i, w)\}.$$

Moreover, for every $p \in Atm$:

$$\begin{aligned} &\mathcal{V}'(p) = \mathcal{V}(p) && \text{if } p \in \mathcal{T}(M) \cup Atm(\varphi), \\ &\mathcal{V}'(p) = \mathcal{N}(i, w) && \text{if } p = f(i, w), \\ &\mathcal{V}'(p) = \emptyset && \text{otherwise.} \end{aligned}$$

It is easy to verify that $\mathcal{N}'(i, w) = \bigcap_{\alpha \in \mathcal{D}'(i,w)} ||\alpha||_{M'}$ for all $i \in Agt$ and for all $w \in W'$ and, more generally, that M' is a finite NDM.

By induction on the structure of φ, We prove that, for all $w \in W$, "$(M, w) \models \varphi$ iff $(M', w) \models \varphi$".

The case $\varphi = p$ is immediate from the definition of $\mathcal{V}'$. By the induction hypothesis, we can prove the boolean cases $\varphi = \neg\psi$ and $\varphi = \psi_1 \wedge \psi_2$ in a straightforward manner.

Let us prove the case $\varphi = \mathtt{E}_i\alpha$.

($\Rightarrow$) Suppose $(M, w) \models \mathtt{E}_i\alpha$. Then, we have $\alpha \in \mathcal{D}(i, w)$. Hence, by the definition of $\mathcal{D}'$, $\alpha \in \mathcal{D}'(i, w)$. Thus, $(M', w) \models \mathtt{E}_i\alpha$.

($\Leftarrow$) Suppose $(M', w) \models \mathtt{E}_i\alpha$. Then, we have $\alpha \in \mathcal{D}'(i, w)$. The definition of $\mathcal{D}'$ ensures that $\alpha \neq f(i, w)$, since $f(i, w) \notin Atm(\mathtt{E}_i\alpha)$. Thus, $\alpha \in \mathcal{D}(i, w)$ and, consequently, $(M, w) \models \mathtt{E}_i\alpha$.

Let us prove the case $\varphi = \mathtt{I}_i\psi$. $(M, w) \models \mathtt{I}_i\psi$ means that $(M, v) \models \psi$ for all $v \in \mathcal{N}(i, w)$. By induction hypothesis and the fact that $\mathcal{N}(i, w) = \mathcal{N}'(i, w)$, the latter is equivalent to $(M', v) \models \psi$ for all $v \in \mathcal{N}'(i, w)$. The latter means that $(M', w) \models \mathtt{I}_i\psi$.

Since M satisfies φ and "$(M, w) \models \varphi$ iff $(M', w) \models \varphi$" for all $w \in W$, M' satisfies φ as well. ■

Equivalence between CMABs and NDMs Our third equivalence result is between CMABs and NDMs.

Theorem 3 *Let $\varphi \in \mathcal{L}$. Then, φ is satisfiable for the class of CMABs if and only if φ is satisfiable for the class of NDMs.*

PROOF. We first prove the left-to-right direction. Let (B, Cxt) be a CMAB with $B = (B_1, \ldots, B_n, V)$ and such that $(B, Cxt) \models \varphi$. We define the structure $M = (W, \mathcal{D}, \mathcal{N}, \mathcal{V})$ as follows:

- $W = \{w_{B'} : B' \in Cxt \cup \{B\}\}$,
- for every $i \in Agt$ and for every $w_{B'} \in W$, if $B' = (\alpha'_{IC}, B'_1, \ldots, B'_n, V')$ then $\mathcal{D}(i, w_{B'}) = B'_i$,
- for every $i \in Agt$ and for every $w_{B'} \in W$, $\mathcal{N}(i, w_{B'}) = \bigcap_{\alpha \in \mathcal{D}(i, w_{B'})} \{w_{B''} \in W : B'' \models \alpha\}$,
- for every $p \in Atm$, $\mathcal{V}(p) = \{w_{B'} \in W : B' \models p\}$.

One can show that M so defined is a NDM. Moreover, by induction on the structure of φ, one can prove that, for all $w_{B'} \in W$, $M, w_{B'} \models \varphi$ iff $(B', Cxt) \models \varphi$. Thus, $(M, w_B) \models \varphi$.

We now prove the right-to-left direction. Let $M = (W, \mathcal{D}, \mathcal{N}, \mathcal{V})$ be a NDM and let w be a world in W such that $(M, w) \models \varphi$. Let us say that a NDM $M = (W, \mathcal{D}, \mathcal{N}, \mathcal{V})$ is non-redundant iff there are no $w, v \in W$ such that $\mathcal{V}^{-1}(w) = \mathcal{V}^{-1}(v)$, and, for all $i \in Agt$, $\mathcal{D}(i, w) = \mathcal{D}(i, v)$. It is straightforward to show that if φ is satisfiable for the class of NDMs then φ is satisfiable for the class of non-redundant NDMs. Thus, from the initial model M, we can find a non-redundant NDM $M' = (W', \mathcal{D}', \mathcal{N}', \mathcal{V}')$ and $v \in W'$ such that $(M', v) \models \varphi$. For every $u \in W'$ we define $B^u = (B^u_1, \ldots, B^u_n, V^u)$ such that $B^u_i = \mathcal{D}(i, u)$ for every $i \in Agt$ and $V^u = \mathcal{V}^{-1}(u)$. Moreover, we define the context $Cxt = \{B^u : u \in W'\}$. One can show that, for every $B^u \in Cxt$, (B^u, Cxt) is a CMAB. The fact that M' is non-redundant is essential to guarantee that there is a one-to-one correspondence between W' and Cxt. By induction on the structure of φ, one can prove that, for all $B^u \in Cxt$, $(B^u, Cxt) \models \varphi$ iff $M', u \models \varphi$. Thus, $B^v \models \varphi$.

■

Axiomatics and decidability

This section is devoted to provide an axiomatization and a decidability result for LDA. To this aim, we first provide a formal definition of this logic.

Definition 9 *We define* LDA *to be the extension of classical propositional logic given by the following axioms and rule of inference:*

$$(\mathtt{I}_i\varphi \wedge \mathtt{I}_i(\varphi \rightarrow \psi)) \rightarrow \mathtt{I}_i\psi \qquad (\mathbf{K}_{\mathtt{I}_i})$$

$$\neg(\mathtt{I}_i\varphi \wedge \mathtt{I}_i\neg\varphi) \qquad (\mathbf{D}_{\mathtt{I}_i})$$

$$\mathtt{E}_i\alpha \rightarrow \mathtt{I}_i\alpha \qquad (\mathbf{Int}_{\mathtt{E}_i,\mathtt{I}_i})$$

$$\frac{\varphi}{\mathtt{I}_i\varphi} \qquad (\mathbf{Nec}_{\mathtt{I}_i})$$

We denote that φ is derivable in LDA by $\vdash_{\mathsf{LDA}} \varphi$. We say that φ is LDA-consistent if $\nvdash_{\mathsf{LDA}} \neg\varphi$.

The logic LDA includes the principles of system KD for the implicit belief operator $\mathtt{I}_i$ as well as an axiom $\mathbf{Int}_{\mathtt{E}_i,\mathtt{I}_i}$ relating explicit belief with implicit belief. Note that there is no consensus in the literature about introspection for implicit belief. For instance, in his seminal work on the logics of knowledge and belief (Hintikka 1962), Hintikka only assumed positive introspection for belief (Axiom 4) and rejected negative introspection (Axiom 5). Other logicians such as (Jones 2015) have argued against the use of both positive and negative introspection axioms for belief. Nonetheless, all approaches unanimously assume that a reasonable notion of implicit belief should satisfy Axioms K and D. In this sense, system KD can be conceived as the minimal logic of implicit belief. On this point, see (Banerjee and Dubois 2014).

To prove our main completeness result, we first prove a theorem about soundness and completeness of LDA for the class of quasi-NDMs.

Soundness and completeness for quasi-NDMs To prove completeness of LDA for the class of quasi-NDMs, we use a canonical model argument.

We consider maximally LDA-consistent sets of formulas in $\mathcal{L}$ (MCSs). The following proposition specifies some usual properties of MCSs.

Proposition 2 *Let Γ be a MCS and let $\varphi, \psi \in \mathcal{L}$. Then:*

- *if $\varphi, \varphi \rightarrow \psi \in \Gamma$ then $\psi \in \Gamma$;*
- *$\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$;*
- *$\varphi \vee \psi \in \Gamma$ iff $\varphi \in \Gamma$ or $\psi \in \Gamma$.*

The following is the Lindenbaum's lemma for our logic. Its proof is standard (cf. Lemma 4.17 in (Blackburn, de Rijke, and Venema 2001)) and we omit it.

Lemma 2 *Let Δ be a LDA-consistent set of formulas. Then, there exists a MCS Γ such that $\Delta \subseteq \Gamma$.*

Let the canonical quasi-NDM model be the tuple $M = (W^c, \mathcal{D}^c, \mathcal{N}^c, \mathcal{V}^c)$ such that:

- W^c is set of all MCSs;
- for all $w \in W^c$, for all $i \in Agt$ and for all $\alpha \in \mathcal{L}_0$, $\alpha \in \mathcal{D}^c(i,w)$ iff $\mathtt{E}_i\alpha \in w$;
- for all $w, v \in W^c$ and for all $i \in Agt$, $v \in \mathcal{N}^c(i,w)$ iff, for all $\varphi \in \mathcal{L}$, if $\mathtt{I}_i\varphi \in w$ then $\varphi \in v$;
- for all $w \in W^c$ and for all $p \in Atm$, $w \in \mathcal{V}^c(p)$ iff $p \in w$.

The next step in the proof consists in stating the following existence lemma. The proof is again standard (cf. Lemma 4.20 in (Blackburn, de Rijke, and Venema 2001)) and we omit it.

Lemma 3 *Let $\varphi \in \mathcal{L}$ and let $w \in W^c$. Then, if $\widehat{\mathtt{I}}_i\varphi \in w$ then there exists $v \in \mathcal{N}^c(i,w)$ such that $\varphi \in v$.*

Then, we prove the following truth lemma.

Lemma 4 *Let $\varphi \in \mathcal{L}$ and let $w \in W^c$. Then, $M^c, w \models \varphi$ iff $\varphi \in w$.*

PROOF. The proof is by induction on the structure of the formula. The cases with φ atomic, Boolean, and of the form $\mathtt{I}_i\psi$ are provable in the standard way by means of Proposition 2 and Lemma 3 (cf. Lemma 4.21 in (Blackburn, de Rijke, and Venema 2001)). The proof for the case $\varphi = \mathtt{E}_i\alpha$ goes as follows: $\mathtt{E}_i\alpha \in w$ iff $\alpha \in \mathcal{D}^c(i,w)$ iff $M^c, w \models \mathtt{E}_i\alpha$. ■

The last step consists in proving that the canonical model belongs to the class **QNDM**.

Proposition 3 *M^c is a quasi-NDM.*

PROOF. Thanks to Axiom ($\mathbf{D}_{\mathtt{I}_i}$), it is easy to prove that M^c satisfies Condition C2 in Definition 7.

Let us prove that it satisfies Condition C1* in Definition 8. To this aim, we just need to prove that if $\alpha \in \mathcal{D}^c(i,w)$ then $\mathcal{N}^c(i,w) \subseteq ||\alpha||_{M^c}$. Suppose $\alpha \in \mathcal{D}^c(i,w)$. Thus, $\mathtt{E}_i\alpha \in w$. Hence, by Axiom ($\mathbf{Int}_{\mathtt{E}_i,\mathtt{I}_i}$) and Proposition 2, $\mathtt{I}_i\alpha \in w$. By the definition of M^c, it follows that, for all $v \in \mathcal{N}^c(i,w)$, $\alpha \in v$. Thus, by Lemma 4, for all $v \in \mathcal{N}^c(i,w)$, $(M^c, v) \models \alpha$. The latter means that $\mathcal{N}^c(i,w) \subseteq ||\alpha||_{M^c}$. ■

The following is our first intermediate result.

Theorem 4 *The logic LDA is sound and complete for the class of quasi-NDMs.*

PROOF. As for soundness, it is routine to check that the axioms of LDA are all valid for the class of quasi-NDMs and that the rule of inference ($\mathbf{Nec}_{\mathtt{I}_i}$) preserves validity.

As for completeness, suppose that φ is a LDA-consistent formula in $\mathcal{L}$. By Lemma 2, there exists $w \in W^c$ such that $\varphi \in w$. Hence, by Lemma 4, there exists $w \in W^c$ such that $M^c, w \models \varphi$. Since, by Proposition 3, M^c is a quasi-NDM, we can conclude that φ is satisfiable for the class of quasi-NDMs. ■

Soundness and completeness for NDMs and CMABs We can state the two main results of this section. The first is about soundness and completeness for the class of NDMs.

Theorem 5 *The logic LDA is sound and complete for the class of NDMs.*

PROOF. It is routine exercise to verify that LDA is sound for the class of NDMs. Now, suppose that formula φ is LDA-consistent. Then, by Theorems 4 and 1, it is satisfiable for the class of finite quasi-NDMs. Hence, by Theorem 2, it is satisfiable for the class of finite NDMs. Thus, more generally, φ is satisfiable for the class of NDMs. ■

The second is about soundness and completeness for the class of CMABs.

Theorem 6 *The logic* LDA *is sound and complete for the class of CMABs.*

SKETCH OF PROOF. The theorem is provable by means of Theorem 5 and Theorem 3. ■

Decidability The second main result of this section is decidability of LDA.

Theorem 7 *The satisfiability problem of* LDA *is decidable.*

PROOF. Suppose φ is satisfiable for the class of NDMs. Thus, by Theorem 5, it is LDA-consistent. Hence, by Theorem 4, it is satisfiable for the class of quasi-NDMs. From the proof of Theorem 1, we can observe that if φ is satisfiable for the class of quasi-NDMs then there exists a quasi-NDM satisfying φ such that (i) its set of worlds contains at most 2^n elements, (ii) the atomic propositions outside $Atm(sub(\varphi))$ are false everywhere in the model, and (iii) the belief base of an agent at a world contains only formulas from $sub(\varphi)$, where n is the size of $sub(\varphi)$. The construction in the proof of Theorem 2 ensures that from this finite quasi-NDM, we can built a finite NDM satisfying φ for which (i) holds and such that (iv) the atomic propositions outside $Atm(sub(\varphi)) \cup X$ are false everywhere in the model, and (v) the belief base of an agent at a world contains only formulas from $sub(\varphi) \cup X$, where X is an arbitrary set of atoms from $Atm \setminus (Atm(\varphi))$ of size at most $2^n \times |Agt|$. Thus, in order to verify whether φ is satisfiable, we fix a X and check satisfiability of φ for all NDMs satisfying (i), (iv) and (v). There are finitely many NDMs of this kind. ■

Related work

The present work lies in the area of logics for non-omniscient agents. Purely syntactic approaches to the logical omniscience problem have been proposed in which an agent's beliefs are described either by a set of formulas which is not necessarily closed under deduction (Eberle 1974; Moore and Hendrix 2011) or by a set of formulas obtained by the application of an incomplete set of deduction rules (Konolige 1986). Logics of time-bounded reasoning have also been studied (Alechina, Logan, and Whitsey 2004; Grant, Kraus, and Perlis 2000), in which reasoning is a represented as a process that requires time due to the time-consuming application of inference rules. Finally, logics of (un)awareness have been studied both in AI (Fagin and Halpern 1987; van Ditmarsch and French 2014; Ågotnes and Alechina 2014) and economics (Modica and Rustichini 1994; Heifetz, Meyer, and Schipper 2006; Halpern and Rêgo 2009).

The closest system to our logic LDA is the logic of local reasoning by (Fagin and Halpern 1987) in which the distinction between explicit and implicit beliefs is also captured. Fagin & Halpern (F&H) use a neighborhood semantics for explicit belief: every agent is associated with a set of sets of worlds, called frames of mind. They define an agent's set of doxastic alternatives as the intersection of the agent's frames of mind. According to F&H's semantics, an agent explicitly believes that φ if and only if she has a frame of mind in which φ is globally true. Moreover, an agent implicitly believes that φ if and only if, φ is true at all her doxastic alternatives. In their semantics, there is no representation of an agent's belief base, corresponding to the set of formulas explicitly believed by the agent. Moreover, differently from our notion of explicit belief, their notion does not completely solve the logical omniscience problem. For instance, while their notion of explicit belief is closed under logical equivalence, our notion is not. Specifically, the following rule of equivalence preserves validity in F&H's logic but not in our logic:

$$\frac{\alpha \leftrightarrow \alpha'}{\mathtt{E}_i\alpha \leftrightarrow \mathtt{E}_i\alpha'}$$

This is a consequence of their use of an extensional semantics for explicit belief. Levesque too provides an extensional semantics for explicit belief with no connection with the notion of belief base (Levesque 1984). In his logic, explicit beliefs are closed under conjunction, while they are not in our logic LDA.

Conclusion

We have presented a logic of explicit and implicit beliefs with a semantics based on belief bases. In the future, we plan to study a variant of this logic in which explicit and implicit beliefs are replaced by *truthful* explicit and implicit knowledge. At the semantic level, we will move from multi-agent belief bases to multi-agent knowledge bases in which the epistemic accessibility relation $\mathcal{R}_i$ is assumed to be reflexive. The logic will include the following extra-axiom:

$$\mathtt{I}_i\varphi \rightarrow \varphi \qquad (\mathbf{T}_{\mathtt{I}_i})$$

We also expect to study a variant of our logic with the following extra-axioms of positive and negative introspection for implicit beliefs:

$$\mathtt{I}_i\varphi \rightarrow \mathtt{I}_i\mathtt{I}_i\varphi \qquad (\mathbf{4}_{\mathtt{I}_i})$$
$$\neg\mathtt{I}_i\varphi \rightarrow \mathtt{I}_i\neg\mathtt{I}_i\varphi \qquad (\mathbf{5}_{\mathtt{I}_i})$$

Moreover, we plan to extend the logic LDA and its epistemic variant by concepts of distributed belief and distribute knowledge.

We will also investigate dynamic extensions of LDA and its epistemic variant by public announcements (Plaza 1989). The core idea is that a public announcement directly affects explicit beliefs. Given the connection between an agent's belief base and her implicit beliefs, it should indirectly affect the latter.

Last but not least, we plan to study the model checking problem for our logic by using the compact representation offered by multi-agent belief models, as defined in Definition 4. As shown by (van Benthem et al. 2015), the possibility of using compact models could be beneficial for model checking.

References

Ågotnes, T., and Alechina, N. 2014. A logic for reasoning about knowledge of unawareness. *Journal of Logic, Language and Information* 23(2):197–217.

Alechina, N.; Logan, B.; and Whitsey, M. 2004. A complete and decidable logic for resource-bounded agents. In *Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*, 606–613. IEEE Computer Society.

Banerjee, M., and Dubois, D. 2014. A simple logic for reasoning about incomplete knowledge. *International Journal of Approximate Reasoning* 55:639–653.

Blackburn, P.; de Rijke, M.; and Venema, Y. 2001. *Modal Logic*. Cambridge: Cambridge University Press.

Dennett, D. C. 1988. Précis of the intentional stance. *Behaviorial and Brain Sciences* 11:495–546.

Dennett, D. C. 1996. *The Intentional Stance*. Cambridge, Massachusetts: MIT Press.

Eberle, R. A. 1974. A logic of believing, knowing and inferring. *Synthese* 26:356–382.

Fagin, R., and Halpern, J. Y. 1987. Belief, awareness, and limited reasoning. *Artificial Intelligence* 34(1):39–76.

Fagin, R.; Halpern, J.; Moses, Y.; and Vardi, M. 1995. *Reasoning about Knowledge*. Cambridge: MIT Press.

Grant, J.; Kraus, S.; and Perlis, D. 2000. A logic for characterizing multiple bounded agents. *Autonomous Agents and Multi-Agent Systems* 3(4):351–387.

Halpern, J. Y., and Rêgo, L. C. 2009. Reasoning about knowledge of unawareness. *Games and Economic Behavior* 67(2):503–525.

Hansson, S. O. 1993. Theory contraction and base contraction unified. *Journal of Symbolic Logic* 58(2):602–625.

Heifetz, A.; Meyer, M.; and Schipper, B. C. 2006. Interactive unawareness. *Journal of Economic Theory* 130:78–94.

Hintikka, J. 1962. *Knowledge and Belief*. New York: Cornell University Press.

Jones, A. J. I. 2015. On the logic of self-deception. *South American Journal of Logic* 1:387–400.

Koieczny, S., and Pérez, R. P. 2002. Merging information under constraints: a logical framework. *Journal of Logic and Computation* 12(5):773–808.

Konolige, K. 1986. *A deduction model of belief*. Los Altos: Morgan Kaufmann Publishers.

Levesque, H. J. 1984. A logic of implicit and explicit belief. In *Proceedings of the Fourth AAAI Conference on Artificial Intelligence (AAAI'84)*, 198–202. AAAI Press.

Lismont, L., and Mongin, P. 1994. On the logic of common belief and common knowledge. *Theory and Decision* 37:75–106.

Lomuscio, A.; Qu, H.; and Raimondi, F. 2015. MCMAS: an open-source model checker for the verification of multi-agent systems. *International Journal on Software Tools for Technology Transfer* 19:1–22.

Makinson, D. 1985. How to give it up: A survey of some formal aspects of the logic of theory change. *Synthese* 62:347–363.

Meyer, J.-J. C., and van der Hoek, W. 1995. *Epistemic Logic for AI and Theoretical Computer Science*. Oxford: Cambridge University Press.

Modica, S., and Rustichini, A. 1994. Awareness and partitional information structures. *Theory and Decision* 37:107–124.

Moore, R. C., and Hendrix, G. G. 2011. Computational models of belief and the semantics of belief sentences. In Peters, S., and Saarinen, E., eds., *Processes, Beliefs, and Questions*, volume 16 of *Synthese Language Library*. Cambridge University Press. 107–127.

Nebel, B. 1992. Syntax-based approaches to belief revision. In Gärdenfors, P., ed., *Belief Revision*. Cambridge: Cambridge University Press. 52–88.

Plaza, J. A. 1989. Logics of public communications. In Emrich, M.; Pfeifer, M.; Hadzikadic, M.; and Ras, Z., eds., *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*.

Rott, A. 1998. "Just because": Taking belief bases seriously. In *Logic Colloquium '98: Proceedings of the 1998 ASL European Summer Meeting*, volume 13 of *Lecture Notes in Logic*, 387–408. Association for Symbolic Logic.

Stalnaker, R. 2002. On the evaluation of solution concepts. *Linguistics and Philosophy* 25(5-6):701–721.

van Benthem, J.; van Eijck, J.; Gattinger, M.; and Su, K. 2015. Symbolic model checking for dynamic epistemic logic. In *Proceedings of the 5th International Workshop on Logic, Rationality and Interaction (LORI 2015)*, volume 9394 of *LNCS*, 366–378. Springer-Verlag.

van Ditmarsch, H., and French, T. 2014. Semantics for knowledge and change of awareness. *Journal of Logic, Language and Information* 23(2):169–195.