24th European Conference on Artificial Intelligence - ECAI 2020
Santiago de Compostela, Spain

A Logic of Explicit and Implicit Distributed Belief

Andreas Herzig[1] and **Emiliano Lorini**[2] and **Elise Perrotin**[3] and **Fabián Romero**[4] and **François Schwarzentruber**[5]

Abstract. We present a new logic of explicit and implicit distributed belief with a formal semantics exploiting the notion of belief base. A coalition's distributed belief of explicit type corresponds to a piece of information contained in the collective belief base of the coalition, which is obtained by pooling together the individual belief bases of its members. A coalition's distributed belief of implicit type corresponds to a piece of information that is derivable from the collective belief base of the coalition. We study axiomatic aspects of our logic as well as complexity of model checking. As distributed belief can be inconsistent (contrary to distributed knowledge), we also study a consistency-preserving variant of distributed belief inspired by the literature on belief merging.

1 Introduction

Epistemic logic (in the broad sense) is the variant of modal logic at the intersection between philosophy [15], artificial intelligence (AI) [9, 26] and economics [19] that is devoted to the formal representation of epistemic attitudes of agents including belief and knowledge. It has been shown to have important applications in AI ranging from security protocols [13, 7] and blockchain protocol [13, 25] to epistemic planning [6] and agent communication protocols [14].

Epistemic logic supports reasoning not only about propositional epistemic attitudes but also about higher-order epistemic attitudes, where the order of an agent's epistemic attitude is defined inductively as follows: (i) an agent's epistemic attitude has order 1 if and only if its content is a propositional formula that does not mention epistemic attitudes of others; (ii) an agent's epistemic attitude has order k with $k > 1$ if and only it is an epistemic attitude about an agent's epistemic attitude of order $k - 1$ (possibly the same agent).

Epistemic logic languages are traditionally interpreted in Kripke's possible worlds semantics. The type of structures used in these semantics are the so-called multi-agent *Kripke models*, namely, multi-relational structures equipped with valuation functions for the interpretation of atomic formulas. Binary relations in a multi-agent Kripke model are called *epistemic accessibility relations*.

Different types of collective attitudes have been defined in the epistemic logic framework and the mathematical properties of their corresponding extensions have been investigated, including axiomatizability and computational complexity. This includes the notions of shared knowledge and belief [11], common knowledge and belief [31, 12], distributed knowledge and belief [12, 32, 1, 28, 30] and collective acceptance [23].

[1] IRIT-CNRS, Toulouse University, France, email: Andreas.Herzig@irit.fr
[2] IRIT-CNRS, Toulouse University, France, email: Emiliano.Lorini@irit.fr
[3] IRIT, Toulouse University, France, email: Elise.Perrotin@irit.fr
[4] IRIT, Toulouse University, France, email: benito.romero@irit.fr
[5] ENS Rennes, France, email: francois.schwarzentruber@ens-rennes.fr

Distributed epistemic attitudes are the results of pooling together or aggregating the agents' individual epistemic attitudes. For instance, for a coalition G of agents to distributively believe that φ, φ must be included in the set of facts that are "aggregatedly" believed by the agents in G. This notion of aggregation is rendered in possible worlds semantics by intersecting the individual epistemic accessibility relations of the coalition's members. However this could be considered too strong as that intersection can easily be empty thereby making distributed belief inconsistent.

The notion of belief aggregation was studied in parallel in the area of belief merging [17, 16, 27]. Nonetheless, there are some fundamental differences between the notion of aggregation studied in epistemic logic, via the notions of distributed knowledge and belief, and the notion of aggregation studied in belief merging. First of all, belief merging focuses exclusively on aggregation of propositional beliefs, while epistemic logic allows us to represent aggregation of higher-order epistemic attitudes. Secondly, the belief merging approach is essentially syntactical, as the aggregation operation is defined on individual belief bases, while the epistemic logic approach is essentially semantical, as aggregation is made at the level of the agents' epistemic accessibility relations.

The aim of this paper is to reconcile the two approaches by presenting a new epistemic logic, with a formal semantics exploiting the notion of belief base like in belief merging, that distinguishes explicit and implicit distributed belief. A coalition's distributed belief of explicit type corresponds to a piece of information contained in the collective belief base of the coalition, which is obtained by pooling together the individual belief bases of its members. A coalition's distributed belief of implicit type corresponds to a piece of information that is derivable from the collective belief base of the coalition. The belief semantics for epistemic logic that we use was introduced in [20] (see also [22, 24, 21]). In this paper, we show that it is well-suited to model aggregation of agents' higher-order individual beliefs. The latter concept is relevant for some AI applications such as recommendation systems, e-democracy, artificial trading agents or autonomous vehicles. For example, in an online recommendation system, it would be interesting to merge not only the agents' individual beliefs about the (good and bad) qualities of a given product, but also the agents' meta-evaluations, namely, what they believe about the others' beliefs about the product qualities.[6] Similarly, in online pools, it could be useful to compute what the agents think about a certain political issue (e.g., efficacy of immigration policies in Europe, increase of the price of gasoline for contrasting climate change, etc.),

[6] In the case of conformity—a well-known phenomenon studied in social psychology [3, 10]—, the agents' beliefs about the product qualities are generally aligned with (and supported by) their higher-order beliefs about the others' beliefs about the product qualities. On the contrary, in the case of anticonformity [33], the agents' 1-order beliefs and their higher-order beliefs about the others' beliefs are generally misaligned.

as well as what they think the other agents think about the same issue. On top of helping with compactness, making a distinction between implicit and explicit belief allows us to separate the aggregation process (at the level of the explicit beliefs) from the deduction process (from explicit to implicit belief), which we believe to be interesting from a conceptual point of view.

The paper is organized as follows. In Section 2, we present the language of our logic of individual and distributed beliefs. Section 3 is devoted to illustrating a semantics for this language which exploits the notion of belief base. In Section 4, we present an axiomatics for the logic, while in Section 5 we study complexity of its model checking problem. Section 6 introduces an introspective variant of the logic in which coalitions are assumed to have introspection over their distributed beliefs. As distributed belief can be inconsistent (contrary to distributed knowledge), in Section 7, we study a consistency-preserving variant of distributed belief inspired by the literature on belief merging. In Section 8, we conclude.

2 A language for distributed doxastic attitudes

This section presents a language for representing agents' individual beliefs and coalitions' distributed beliefs of both explicit and implicit type. It extends the language of individual explicit and implicit belief presented in [20] with distributed belief. Assume a countably infinite set of atomic propositions $Atm = \{p, q, \dots\}$ and a finite set of agents $Agt = \{1, \dots, n\}$. We define the language in two steps.

We first define the language $\mathcal{L}_0(Atm, Agt)$ by the following grammar in Backus-Naur Form (BNF):

$$\alpha \quad ::= \quad p \mid \neg\alpha \mid \alpha_1 \wedge \alpha_2 \mid \triangle_i \alpha,$$

where p ranges over Atm and i ranges over Agt. $\mathcal{L}_0(Atm, Agt)$ is the language for representing distributed explicit beliefs of coalitions. The formula $\triangle_i \alpha$ is read "agent i explicitly believes that α".

The language $\mathcal{L}_1(Atm, Agt)$ extends the language $\mathcal{L}_0(Atm, Agt)$ by modal operators of distributed implicit belief and is defined by the following grammar:

$$\varphi \quad ::= \quad \alpha \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_G \varphi,$$

where α ranges over $\mathcal{L}_0(Atm, Agt)$ and G ranges over $2^{Agt*} = 2^{Agt} \setminus \{\emptyset\}$. For notational convenience we write $\mathcal{L}_0$ instead of $\mathcal{L}_0(Atm, Agt)$ and $\mathcal{L}_1$ instead of $\mathcal{L}_1(Atm, Agt)$, 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.

The formula $\Box_G \varphi$ is read "coalition G has the distributed implicit belief that φ". We define the dual operator $\Diamond_G$ as follows:

$$\Diamond_G \varphi \overset{\text{def}}{=} \neg\Box_G\neg\varphi.$$

$\Diamond_G \varphi$ has to be read "φ is compatible (or consistent) with coalition G's explicit beliefs".

For notational convenience, for every $i \in Agt$, we simply write $\Box_i \varphi$ instead of $\Box_{\{i\}} \varphi$ to represent agent i's implicit belief that φ.

3 Belief base semantics

Following [20], in this section, we present a formal semantics for the language $\mathcal{L}_1$ exploiting belief bases. Unlike the standard Kripke semantics for epistemic logic in which the notions of possible world (or state) and epistemic alternative are given as primitive, in this semantics they are are defined from the primitive concept of belief base.

Definition 1 (State) *A state is a tuple $B = (B_1, \dots, B_n, V)$ 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.*

The set of all states is denoted by $\mathbf{S}$.

The sublanguage $\mathcal{L}_0(Atm, Agt)$ is interpreted with respect to states, as follows.

Definition 2 (Satisfaction relation) *Let $B = (B_1, \dots, B_n, V) \in \mathbf{S}$. Then:*

$$\begin{aligned}
B \models p &\iff p \in V,\\
B \models \neg\alpha &\iff B \not\models \alpha,\\
B \models \alpha_1 \wedge \alpha_2, &\iff B \models \alpha_1 \text{ and } B \models \alpha_2,\\
B \models \triangle_i \alpha &\iff \alpha \in B_i.
\end{aligned}$$

Observe in particular the set-theoretic interpretation of the distributed explicit belief operator: agent i explicitly believes that α if and only if α is included in her belief base.

It is also worth considering belief correct states, according to which every fact that an agent explicitly believes has to be true.

Definition 3 (Belief correct state) *The state $B = (B_1, \dots, B_n, V)$ is said to be belief correct if and only if, for every $i \in Agt$ and for every $\alpha \in B_i$, if $\alpha \in B_i$ then $B \models \alpha$. The set of all belief correct states is denoted by $\mathbf{S}_{BC}$.*

The following definition introduces the notion of "pooling", a simple form of belief base aggregation which consists in taking the union of the belief bases of the agents in the coalition.

Definition 4 (Pooling) *Let $B = (B_1, \dots, B_n, V) \in \mathbf{S}$ and let $G \in 2^{Agt*}$. Then*

$$Pool_G(B) = \bigcup_{i \in G} B_i.$$

A multi-agent belief model (MAB) is defined to be a state supplemented with a set of states, called *context*. The latter includes all states that are compatible with the common ground [29], i.e., the body of information that the agents commonly believe to be the case.

Definition 5 (Multi-agent belief model) *A multi-agent belief model (MAB) is a pair (B, Cxt), where $B \in \mathbf{S}$ and $Cxt \subseteq \mathbf{S}$. The class of MABs is denoted by $\mathbf{MAB}$.*

Note that in this definition we do not require $B \in Cxt$. A MAB (B, Cxt) such that $Cxt = \mathbf{S}$ is said to be *complete*, since $\mathbf{S}$ is conceivable as the complete (or universal) context which contains all possible states. The following definition introduces the concept of doxastic alternative.

Definition 6 (Doxastic alternatives) *Let $G \in 2^{Agt*}$. Then $\mathcal{R}_G$ is the binary relation on the set of states $\mathbf{S}$ such that, for all $B = (B_1, \dots, B_n, V), B' = (B'_1, \dots, B'_n, V') \in \mathbf{S}$:*

$$B \mathcal{R}_G B' \text{ if and only if } \forall \alpha \in Pool_G(B) : B' \models \alpha.$$

$B \mathcal{R}_G B'$ means that B' is a doxastic alternative for coalition G at B. The idea of the previous definition is that B' is a doxastic alternative for coalition G at B if and only if, B' satisfies all facts that are in the database resulting from pooling together the belief bases of the agents in the coalition.

The collective doxastic relation computed from the collective belief base is equal to the intersection of the individual doxastic relations computed from the individual belief bases:

Proposition 1 *Let $G \in 2^{Agt*}$. Then $\mathcal{R}_G = \bigcap_{i \in G} \mathcal{R}_{\{i\}}$.*

The following definition extends Definition 2 to the full language $\mathcal{L}_1$. Its formulas are interpreted with respect to MABs. (We omit Boolean cases, as they are defined in the usual way.)

Definition 7 (Satisfaction relation (cont.)) *Let (B, Cxt) be a MAB. Then:*

$$(B, Cxt) \models \alpha \iff B \models \alpha,$$
$$(B, Cxt) \models \Box_G \varphi \iff \text{for all } B' \in Cxt : \text{if } B \mathcal{R}_G B' \text{ then } (B', Cxt) \models \varphi.$$

We consider the subclass of MABs that guarantee correctness of the agents' beliefs.

Definition 8 (Belief correct MAB) *The MAB (B, Cxt) is belief correct (BC) if and only if $B \in Cxt$ and, for every $i \in Agt$ and for every $B' \in Cxt$, $B' \mathcal{R}_i B'$. The class of MABs satisfying BC is denoted by $\mathbf{MAB}_{BC}$.*

Saying that (B, Cxt) satisfies BC is the same thing as saying that $B \in Cxt$ and, for every $i \in Agt$, the relation $\mathcal{R}_i \cap (Cxt \times Cxt)$ is reflexive. The condition $B \in Cxt$ in Definition 8 is necessary to make the agents' implicit beliefs correct, i.e., to make the formula $\Box_i \varphi \rightarrow \varphi$ valid. For example, for $B = (\{p\}, \{p \rightarrow q\}, \{p, q\})$ we have that both $(B, \{B\})$ and $(B, \{B, (\emptyset, \emptyset, \{p, q\})\})$ satisfy BC.

As the following proposition highlights, belief correctness for MABs is completely characterized by the fact that the actual world is included in the agents' common ground and that the agents' explicit beliefs are correct in the sense of Definition 3.

Proposition 2 *A MAB (B, Cxt) satisfies BC if and only if $B \in Cxt$ and $Cxt \subseteq \mathbf{S}_{BC}$.*

Let $\varphi \in \mathcal{L}_1$. We say that φ is valid for the class $\mathbf{MAB}$ (resp. $\mathbf{MAB}_{BC}$) if and only if, for every $(B, Cxt) \in \mathbf{MAB}$ (resp. $(B, Cxt) \in \mathbf{MAB}_{BC}$) we have $(B, Cxt) \models \varphi$. We say that φ is satisfiable for the class $\mathbf{MAB}$ (resp. $\mathbf{MAB}_{BC}$) if and only if $\neg\varphi$ is not valid for the class $\mathbf{MAB}$ (resp. $\mathbf{MAB}_{BC}$).

Before concluding this section let us illustrate the belief base semantics for distributed explicit and implicit belief in a concrete example.

Example *Suppose Agt is a community of autonomous vehicles that have to get from a point A to a point B. There are two routes from A to B, r_0 and r_1. Route r_0 is the shortest one so that each vehicle prefers to take it rather than to take route r_1, if it does not believe that the others will also take it. Indeed, if all vehicles take the same route, then traffic will be congested and every vehicle will waste a considerable amount of time. The propositional variable $take_{i,r_x}$ denotes that agent i takes route x, for $x \in \{0, 1\}$. Let us assume that each vehicle has the following pieces of information in its belief base:*

- *if a vehicle does not believe explicitly that the other vehicles will take route r_0, then it will take route r_0,*
- *and if all vehicles take route r_0 (resp. r_1), then the traffic in r_0 (resp. r_1) will be congested.*

Moreover, let us assume that each vehicle:

- *explicitly believes that it does not believe explicitly that the other vehicles will take route r_0.*

In formal terms, for every $i \in Agt$ we assume:

$$B_i = \{\beta, \gamma, \delta_i\}$$

where

$$\beta \stackrel{\text{def}}{=} \bigwedge_{i \in Agt, x \in \{0,1\}} \big(\neg \triangle_i take_{Agt \setminus \{i\}, r_0} \rightarrow take_{i, r_0}\big),$$
$$\gamma \stackrel{\text{def}}{=} \bigwedge_{x \in \{0,1\}} (take_{Agt, r_x} \rightarrow congested_{r_x}),$$
$$\delta_i \stackrel{\text{def}}{=} \neg \triangle_i take_{Agt \setminus \{i\}, r_0},$$

and, for every $G \in 2^{Agt*}$ and for every $x \in \{0, 1\}$:

$$take_{G, r_x} \stackrel{\text{def}}{=} \bigwedge_{i \in G} take_{i, r_x}.$$

Note that the explicit belief that β is an agent's explicit belief of order 2 since it mentions the agents' explicit beliefs of order 1. It is a routine exercise to verify that, for every $B = (B_1, \ldots, B_n, V) \in \mathbf{S}$ where V is an arbitrary valuation, we have:

$$(B, \mathbf{S}) \models \bigwedge_{i \in Agt} \neg \Box_i congested_{r_0} \wedge \Box_{Agt} congested_{r_0}.$$

This means that, while none of the vehicles has the implicit belief that route r_0 will be congested, the community of vehicles has the distributed implicit belief that route r_0 will be congested.

4 Axiomatics

We define the logic LDA_n^D axiomatized by the following principles:

Propositional calculus	(PC)
$\big(\Box_i \varphi \wedge \Box_i(\varphi \rightarrow \psi)\big) \rightarrow \Box_i \psi$	$(K_{\Box_i})$
$\triangle_i \alpha \rightarrow \Box_i \alpha$	$(Int_{\triangle_i, \Box_i})$
$\dfrac{\varphi}{\Box_i \varphi}$	$(Nec_{\Box_i})$
$\Box_G \varphi \rightarrow \Box_{G'} \varphi$ if $G \subseteq G'$	$(Mon_{\Box_G})$

where $Agt = \{1, \ldots, n\}$, $i \in Agt$ and $G, G' \subseteq Agt$.

LDA stands for "Logic of Doxastic Attitudes", an acronym introduced in [20] (see also [22]). For the axioms we take inspiration from that same paper as well as [9]. We also define $\mathsf{LDA}_n^{D,T}$ by adding to the above the following axiom:

$$\Box_{Agt} \varphi \rightarrow \varphi \qquad (T_{\Box_{Agt}})$$

Soundness of LDA_n^D relative to $\mathbf{MAB}$ and of $\mathsf{LDA}_n^{D,T}$ relative to $\mathbf{MAB}_{BC}$ is straightforward. To prove completeness of the first item, we first show completeness relative to another class of models, namely notional doxastic models with distributed beliefs (NDMDs). The class of notional doxastic models (NDMs) was introduced for the first time in [20].[7]

To do this we start from completeness relative to a weaker class of models called quasi-NDMDs, and then we show how to apply several transformations to a quasi-NDMD satisfying a given formula φ in order to turn it into a NDMD satisfying φ. Completeness of the second item follows by remarking that a reflexivity property is preserved throughout all of these transformations. For the sake of brevity we will omit finer details of the proofs.

[7] The term 'notional' is borrowed from [8] (see, also, [18]). According to Dennett, an agent's notional world is a world at which all the agent's explicit beliefs are true.

Definition 9 (NDMD) *A notional doxastic model with distributed beliefs is a tuple* $M = (W, B, R, V)$ *where* W *is a set of worlds,* $B : Agt \times W \longrightarrow 2^{\mathcal{L}_0}$ *is a doxastic function,* $R : 2^{Agt} \times W \longrightarrow 2^W$ *is a notional function and* $V : Atm \longrightarrow 2^W$ *is a valuation function, and, following the semantics:*

$$\begin{aligned}
(M,w) \models p &\iff w \in V(p),\\
(M,w) \models \neg\varphi &\iff (M,w) \not\models \varphi,\\
(M,w) \models \varphi \wedge \psi &\iff (M,w) \models \varphi \text{ and } (M,w) \models \psi,\\
(M,w) \models \triangle_i \alpha &\iff \alpha \in B(i,w),\\
(M,w) \models \Box_G \alpha &\iff \forall v \in R(G,w) : (M,v) \models \varphi.
\end{aligned}$$

M satisfies the following conditions:

$$R(\{i\}, w) = \bigcap_{\alpha \in B(i,w)} \|\alpha\|_M \qquad (\mathsf{NDMD}_1)$$

$$R(G, w) = \bigcap_{i \in G} R(\{i\}, w) \qquad (\mathsf{NDMD}_2)$$

where $\|\alpha\|_M$ *is the truth set of* α *in* M*. If* $w \in R(Agt, w)$ *for all* $w \in W$ *then we say that* M *is reflexive.*

Definition 10 (Quasi-NDMD) *A quasi-NDMD is a model* $M = (W, B, R, V)$ *following the same definition as above except conditions* (NDMD_1) *and* (NDMD_2) *are replaced by the following:*

$$R(\{i\}, w) \subseteq \bigcap_{\alpha \in B(i,w)} \|\alpha\|_M \qquad (\mathsf{QNDMD}_1)$$

$$R(G, w) \subseteq \bigcap_{i \in G} R(\{i\}, w) \qquad (\mathsf{QNDMD}_2)$$

Proposition 3 *If a given formula* φ *is consistent in* LDA_n^D *(resp.* $\mathsf{LDA}_n^{D,T}$*) then it is satisfiable in a quasi-NDMD (resp. a reflexive quasi-NDMD).*

PROOF. If φ is consistent in LDA_n^D then it is satisfiable in the canonical model $M^C = (W^C, B^C, R^C, V^C)$, where:

- W^C is the set of maximal consistent sets for LDA_n^D;
- for all $w \in W^C$, $i \in Agt$, and $\alpha \in \mathcal{L}_0$, $\alpha \in B^C(i, w)$ iff $\triangle_i \alpha \in w$;
- for all $w, v \in W^C$ and $G \subseteq Agt$, $v \in R(G, w)$ iff for all φ such that $\Box_G \varphi \in w$, $\varphi \in v$;
- for all $w \in W^C$ and $p \in Atm$, $w \in V^C(p)$ iff $p \in w$.

M^C is a quasi-NDMD. If φ is consistent in $\mathsf{LDA}_n^{D,T}$ then we can define in the same manner a canonical model M_T^C in which worlds are the maximal consistent sets for $\mathsf{LDA}_n^{D,T}$. M_T^C is a reflexive quasi-NDMD. ■

The next step is to recover (NDMD_2). For this we first use a filtration argument so that we can work with finite quasi-NDMDs.

Lemma 1 *If a formula* φ *is satisfiable in a (reflexive) quasi-NDMD then it is satisfiable in a finite (reflexive) quasi-NDMD.*

SKETCH OF PROOF. If $M = (W, B, R, V)$ is a quasi-NDMD and Σ is a finite set of formulas closed under subformulas, define $M_\Sigma = (W_\Sigma, B_\Sigma, R_\Sigma, V_\Sigma)$ such that:

- $W_\Sigma = \{[w]_\Sigma : w \in W\}$, where $[w]_\Sigma = \{v \in W : \forall \varphi \in \Sigma, (M, w) \models \varphi \text{ iff } (M, v) \models \varphi\}$;
- for all $i \in Agt$ and $[w]_\Sigma \in W_\Sigma$, $B_\Sigma(i, [w]_\Sigma) = \Sigma \cap \bigcap_{w' \in [w]_\Sigma} B(i, w')$;
- for all $G \subseteq Agt$ and $[w]_\Sigma \in W_\Sigma$, $R_\Sigma(G, [w]_\Sigma) = \{[v]_\Sigma : \exists w' \in [w]_\Sigma, \exists v' \in [v]_\Sigma, v' \in R(G, w')\}$;
- for all $p \in Atm$, $V_\Sigma(p) = \{[w]_\Sigma : (M, w) \models p\}$ if $p \in Atm(\Sigma)$ and $V'(p) = \emptyset$ otherwise.

Then M_Σ is a finite quasi-NDMD such that for any $w \in W$ and any formula $\psi \in \Sigma$, $(M, w) \models \psi$ iff $(M_\Sigma, [w]_\Sigma) \models \psi$. By taking $\Sigma = sub(\varphi)$, where $sub(\varphi)$ is the set of subformulas of φ, we obtain the desired model. Moreover if M is reflexive then M_Σ is also reflexive. ■

Lemma 2 *For any quasi-NDMD* $M = (W, B, R, V)$ *there is an equivalent model* $M' = (W, B, R', V)$ *such that for all* $w \in W$ *and* $G, G' \subseteq Agt$*, if* $G \subseteq G'$ *then* $R'(G', w) \subseteq R'(G, w)$*. Our meaning of equivalent is: for any formula* φ *and* $w \in W$*,* $(M, w) \models \varphi$ *iff* $(M', w) \models \varphi$*, and if* M *is reflexive then* M' *is also reflexive. We call models satisfying this property of* R downwards-closed*.*

If $M = (W, B, R, V)$ is a finite downwards-closed quasi-NDMD where $W = \{w_1, \ldots, w_n\}$, define the expansion M' of M as follows: $M' = (W', B', R', V')$ where

- $W' = \bigcup_{i \leqslant n} W_i$, where for all $i \leqslant n$, $W_i = \{z_\sigma : \sigma \in S_{1,i} \times \cdots \times S_{n,i}\}$, and for all $i, k \leqslant n$, $S_{k,i} = \{G \subseteq Agt : w_i \in R(G, w_k) \text{ and } \forall G' \subseteq Agt\}$, if $G \subsetneq G'$ then $w_i \notin R(G', w_k)$;
- for all $i \leqslant n$, $w \in W_i$ and $j \in Agt$, $B'(j, w) = B(j, w_i)$;
- for all $i, j \leqslant n$, $w \in W_i$, $v = z_\sigma \in W_j$ and $G \subseteq Agt$, $v \in R'(G, w)$ iff $G \subseteq \sigma[i]$;
- for all $p \in Atm$, $i \leqslant n$ and $w \in W_i$, $w \in V'(p)$ iff $w_i \in V(p)$.

Lemma 3 *M' is a finite quasi-NDMD verifying* (NDMD_2)*. Moreover for any formula* φ*,* $i \leqslant n$ *and* $w \in W_i$*,* $(M', w) \models \varphi$ *iff* $(M, w_i) \models \varphi$*, and if* M *is reflexive then* M' *is also reflexive.*

Proposition 4 *If* φ *is satisfiable in a (reflexive) quasi-NDMD then there exists a finite (reflexive) quasi-NDMD* $M = (W, B, R, V)$ *satisfying* φ *and verifying* (NDMD_2)*.*

SKETCH OF PROOF. If φ is satisfiable in a (reflexive) quasi-NDMD then φ is also satisfiable in a finite (reflexive) quasi-NDMD, and in a finite (reflexive) downwards-closed quasi-NDMD. By expanding this last model in the manner described above we get a model satisfying φ which verifies the desired properties. ■

We can now recover (NDMD_1) by expanding belief bases in order to contract $\bigcap_{\alpha \in B(i,w)} \|\alpha\|_M$ for a given i and w in a given model M.

Proposition 5 *If* φ *is satisfiable in a finite (reflexive) quasi-NDMD* $M = (W, B, R, V)$ *verifying* (NDMD_2)*, then* φ *is satisfiable in a finite (reflexive) NDMD.*

SKETCH OF PROOF. Let $\mathcal{T}(M) = \bigcup_{w \in W, i \in Agt} Atm(B(i, w))$ be the (finite) terminology of M and let $f : Agt \times W \longrightarrow Atm \setminus (\mathcal{T}(M) \cup Atm(\varphi)$ be an injective function (such a function exists because Atm is infinite). Define $M' = (W, B', R, V')$ such that for all $i \in Agt$ and $w \in W$, $B'(i, w) = B(i, w) \cup \{f(i, w)\}$ and

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

We can show by induction on the structure of ψ that for all $\psi \in sub(\varphi)$ and for all $w \in W$, $(M, w) \models \psi$ iff $(M', w) \models \psi$. In particular $(M', w) \models \varphi$. Moreover M' is a finite NDMD, and if M is reflexive then M' is still reflexive. ■

Proposition 6 *If φ is satisfiable in a finite NDMD (resp. a finite reflexive NDMD) then φ is satisfiable in $\mathbf{MAB}$ (resp. $\mathbf{MAB}_{BC}$).*

SKETCH OF PROOF. If $Agt = \{1, \ldots, n\}$, let $M = (W, B, R, V)$ be a finite NDMD satisfying φ. Define for each $w \in W$, $B^w = (B^w_1, \ldots, B^w_n, V^w)$ where $V^w = \{p \in Atm : w \in V(p)\}$ and for all $i \in Agt$, $B^w_i = B(i, w)$. Define $Cxt = \{B^w : w \in W\}$. Then for any formula ψ and any $w \in W$, $(M, w) \models \psi$ iff $(B^w, Cxt) \models \psi$, and in particular if $(M, w) \models \varphi$ for some $w \in W$ then $(B^w, Cxt) \models \varphi$. Moreover if M is reflexive then $(T_{\Box_{Agt}})$ is valid in M and therefore in Cxt, hence (B^w, Cxt) satisfies BC. ■

Corollary 1 *If a formula φ is consistent in LDA^D_n (resp. $\mathsf{LDA}^{D,T}_n$) then it is satisfiable in $\mathbf{MAB}$ (resp. $\mathbf{MAB}_{BC}$).*

5 Model checking

The following is a compact formulation of the model checking problem for the language $\mathcal{L}_1$.

Model checking
Given: $\varphi \in \mathcal{L}_1$ and a finite $B \in \mathbf{S}$.
Question: Do we have $(B, \mathbf{S}) \models \varphi$?

Belief correct model checking
Given: $\varphi \in \mathcal{L}_1$ and a finite $B \in \mathbf{S}_{BC}$.
Question: Do we have $(B, \mathbf{S}_{BC}) \models \varphi$?

where the state $B = (B_1, \ldots, B_n, V)$ is said to be finite if V and every B_i are finite.

Note that, thanks to Proposition 2, the MAB $(B, \mathbf{S}_{BC})$ in the belief correct variant of model checking belongs to the model class $\mathbf{MAB}_{BC}$, as expected.

In [21], it is proved that the previous two problems are PSPACE-hard, when considering the fragment of $\mathcal{L}_1$ with only singleton coalition operators of type $\Box_{\{i\}}$. We are going to prove that the two problems are in PSPACE.

Figure 1 shows a generic algorithm for model checking of a formula φ in a given finite state B with an abstract function rel. With the logic defined in this section, the call $rel(B, B', G)$ checks that $B\mathcal{R}_G B'$. Since α in checking that $B' \models \alpha$ (see Definition 6) does not contain any implicit belief operator, it is reducible to the propositional problem by stating any explicit belief as a fresh proposition, hence, checking $B\mathcal{R}_G B'$ can be performed in polynomial time and space.

The following theorem establishes PSPACE membership of the model checking problem.

Theorem 1 *If rel is evaluated in polynomial space, then the generic model checking algorithm of Figure 1 runs in polynomial space.*

SKETCH OF PROOF. The depth of nested calls in $mc(B, \varphi)$ is bounded by the size of φ. The local memory used by the recursive call is polynomial in the size of the initial B and the size of φ. Note that the loop "**for** all $B' \ldots$" can be performed by enumerating the B' containing correct subformulas of formulas in the initial B and in the initial formula φ. There is an exponential number of such B' but storing the current B' requires only a polynomial amount of space.

For belief correct model checking we just check each time that the states B' under consideration are correct: this can be done in polynomial time, hence in polynomial space. ■

```
procedure mc(B, φ)
    match φ do
        case p: return B ⊨ p
        case ¬ψ: return not mc(B, ψ)
        case ψ1 ∧ ψ2: return mc(B, ψ1) and mc(B, ψ2)
        case △_i α: return α ∈ B_i
        case □_G ψ:
            for all B' such that rel(B, B', G) do
                if not mc(B', ψ) return false
            return true
```

Figure 1. Generic algorithm for model checking.

Note that the algorithm for belief correct model checking is the same, except that we check that each state considered during the execution is belief correct (Definition 3).

Corollary 2 *Both model checking and belief correct model checking are in PSPACE.*

The algorithm of Figure 1 and Theorem 1 are important for the other PSPACE-membership results given in the rest of the paper.

6 Introspective variant

The following definition introduces a variant of the notion of doxastic alternative for potentially introspective (pi) coalitions.

Definition 11 (Doxastic alternatives for pi coalitions) *Let $G \in 2^{Agt*}$. Then, $\mathcal{R}^{pi}_G$ is the binary relation on the set $\mathbf{S}$ such that, for all $B = (B_1, \ldots, B_n, V), B' = (B'_1, \ldots, B'_n, V') \in \mathbf{S}$:*

$$B\mathcal{R}^{pi}_G B' \text{ if and only if}$$
$$(i)\ \forall \alpha \in Pool_G(B) : B' \models \alpha, \text{ and}$$
$$(ii)\ Pool_G(B) = Pool_G(B').$$

$B\mathcal{R}^{pi}_G B'$ means that B' is a doxastic alternative for the potentially introspective coalition G at B. According to the previous definition, if coalition G is potentially introspectively then her set of doxastic alternatives at B (i.e., $\mathcal{R}^{pi}_G(B)$) includes all and only those states that satisfy the coalition's explicit beliefs, and that are for coalition G subjectively equivalent to B. Note that $\mathcal{R}^{pi}_G \subseteq \mathcal{R}_G$ since item (i) is exactly the definition of the relation $\mathcal{R}_G$.

As the following proposition indicates, if a coalition is potentially introspective, then its implicit beliefs are closed under positive and negative introspection.

Proposition 7 *Let $G \in 2^{Agt*}$. Then, the relation $\mathcal{R}^{pi}_G$ is transitive and Euclidean.*

PROOF. We first prove transitivity. Suppose $B\mathcal{R}^{pi}_G B'$ and $B'\mathcal{R}^{pi}_G B''$. The latter implies that $B' \models \alpha$ for all $\alpha \in Pool_G(B)$, $B'' \models \alpha$ for all $\alpha \in Pool_G(B')$ and $Pool_G(B) = Pool_G(B') = Pool_G(B'')$. Hence, $B'' \models \alpha$ for all $\alpha \in Pool_G(B)$, since $Pool_G(B) = Pool_G(B')$. It follows that $B\mathcal{R}_G^{pi}B''$.

Let us prove that $\mathcal{R}_G^{pi}$ is Euclidean. Suppose $B\mathcal{R}_G^{pi}B'$ and $B\mathcal{R}_G^{pi}B''$. The latter implies that $B' \models \alpha$ and $B'' \models \alpha$ for all $\alpha \in Pool_G(B)$, and $Pool_G(B) = Pool_G(B') = Pool_G(B'')$. Hence, $B'' \models \alpha$ for all $\alpha \in Pool_G(B')$, since $Pool_G(B) = Pool_G(B')$. It follows that $B'\mathcal{R}_G^{pi}B''$. ■

The following proposition generalizes Proposition 1 to the introspective case.

Proposition 8 *Let $G \in 2^{Agt*}$. Then, $\mathcal{R}_G^{pi} = \bigcap_{i\in G} \mathcal{R}_{\{i\}}^{pi}$.*

By means of the new relation $\mathcal{R}_G^{pi}$, we can define a variant of the language $\mathcal{L}_1(Atm, Agt)$, denoted by $\mathcal{L}_2(Atm, Agt)$, in which every implicit belief operator $\Box_G$ is replaced by a potentially introspective variant of it of the form $\Box_G^{pi}$, where $\Box_G^{pi}\varphi$ has to be read "if coalition G was potentially introspective, it would implicitly believe that φ" (or "coalition G implicitly believes that φ, under the assumption that it is potentially introspectively"). The language $\mathcal{L}_2(Atm, Agt)$ is defined by the following grammar:

$$\varphi \quad ::= \quad \alpha \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_G^{pi}\varphi,$$

where α ranges over $\mathcal{L}_0(Atm, Agt)$ and G ranges over 2^{Agt*}.

Like the operator $\Box_G$, the operator $\Box_G^{pi}$ is interpreted relative to a MAB (B, Cxt), as follows:

$$(B, Cxt) \models \Box_G^{pi}\varphi \iff \forall B' \in Cxt : \text{if } B\mathcal{R}_G^{pi}B' \text{ then } (B', Cxt) \models \varphi.$$

Model checking and belief correct model checking for $\mathcal{L}_2$ are the same as the model checking problems defined in Section 5 except that the input formula φ is now in $\mathcal{L}_2$.

Theorem 2 *Model checking and belief correct model checking for $\mathcal{L}_2$ is in PSPACE.*

SKETCH OF PROOF. To satisfy the hypothesis of Theorem 1, we need to check that $B\mathcal{R}_G^{pi}B'$ can be checked in polynomial space. Note that formulas α are in $\mathcal{L}_0$ in Point (i) in Definition 11. So checking that $B' \models \alpha$ can be performed in polynomial time, since it is close to model checking of a propositional formula. Therefore $B\mathcal{R}_G^{pi}B'$ can be checked in polynomial time, and thus in polynomial space. ■

7 Consistency

The pooling operation we defined in Section 3 does not necessarily guarantee consistency of the collective belief base resulting from the aggregation of the individual belief bases. Indeed, two agents in a coalition may have contradictory explicit beliefs so that, by pooling together their belief bases, we obtain an inconsistent belief base for the collective. In this section we study a different belief aggregation operation, called belief base combination, that warrants belief base consistency. Before defining it, we need some preliminary notions.

We denote by $MCS_G(B)$ the set of maximally consistent subsets of coalition G's belief base.

Definition 12 *Let $G \in 2^{Agt*}$ and $B \in \mathbf{S}$. Then, $X \in MCS_G(B)$ if and only if:*

- *$X \subseteq Pool_G(B)$,*
- *$||X||_{\mathbf{S}} \neq \emptyset$, and*
- *there is no $X' \subseteq Pool_G(B)$ such that $X \subset X'$ and $||X'||_{\mathbf{S}} \neq \emptyset$,*

where, for every $Y \subseteq \mathcal{L}_0$, $||Y||_{\mathbf{S}} = \{B' \in \mathbf{S} : \forall\alpha \in Y, B' \models \alpha\}$.

It is worth noting that the definition of $MCS_G(B)$ can be formulated in terms of propositional consistency. To see this, let $\mathcal{L}_{\mathsf{PROP}}$ be the propositional language built from the set of atomic formulas $Atm^+ = Atm \cup \{p_{i,\alpha} : i \in Agt \text{ and } \alpha \in \mathcal{L}_0\}$ and let tr be the following translation from $\mathcal{L}_0$ to $\mathcal{L}_{\mathsf{PROP}}$:

$$tr(p) = p \qquad tr(\neg\alpha) = \neg tr(\alpha),$$
$$tr(\alpha_1 \wedge \alpha_2) = tr(\alpha_1) \wedge tr(\alpha_2), \qquad tr(\triangle_i\alpha) = p_{i,\alpha}.$$

For each $X \subseteq \mathcal{L}_0$, we define $tr(X) = \{tr(\alpha) : \alpha \in X\}$. Moreover, we say that X is propositionally consistent if and only if $\bot \notin Cn\big(tr(X)\big)$, where Cn is the classical deductive closure operator over the propositional language $\mathcal{L}_{\mathsf{PROP}}$. The following holds.

Proposition 9 *Let $X \subseteq \mathcal{L}_0$. Then, $||X||_{\mathbf{S}} \neq \emptyset$ if and only if $tr(X)$ is propositionally consistent.*

PROOF. Let W be the set of all valuations for the propositional language $\mathcal{L}_{\mathsf{PROP}}$. There exists a bijection $f : W \longrightarrow \mathbf{S}$ such that $(f(w), \mathbf{S}) \models \alpha$ iff $w \models tr(\alpha)$, for all $w \in W$ and $\alpha \in \mathcal{L}_0$. Propositional consistency of $tr(X)$ means that we can find a valuation $w \in W$ where all formulas in $tr\big(X\big)$ hold. Thus, there exists $B' \in \mathbf{S}$ such that $f(w) = B'$ and $(B', \mathbf{S}) \models \alpha$ for all $\alpha \in X$. The left-to-right direction can be proved in an analogous way. ■

By the previous proposition, the set $MCS_G(B)$ can be defined in an equivalent way by replacing $||X||_{\mathbf{S}} \neq \emptyset$ by $\bot \notin Cn\big(tr(X)\big)$ in the first item and $||X'||_{\mathbf{S}} \neq \emptyset$ by $\bot \notin Cn\big(tr(X')\big)$ in the second item of Definition 12.

The following definition introduces the notion of belief base combination.

Definition 13 (Combining) *Let $B = (B_1, \ldots, B_n, V) \in \mathbf{S}$. Then,*

$$Comb_G(B) = \bigcap_{X \in MCS_G(B)} X.$$

$Comb_G(B)$ is the result of combining the individual belief bases in B of coalition G's members. This notion of belief base combination is inspired by [4, 5]. It is slightly different from belief merging in the sense of [17] since it takes the intersection of *all* maximally consistent subsets of coalition G's belief base, while in belief merging the intersection of a *selection* of the MCSs is taken. This parallels the distinction between full meet contraction and partial meet contraction in the belief revision literature [2].[8] We leave for future work the analysis of a notion of distributed belief more in line with such a notion of belief merging using a selection function on the set $MCS_G(B)$.

Doxastically consistent (dc) coalitions are coalitions whose doxastic accessibility relations are computed after having combined the individual belief bases of its members.

Definition 14 (Doxastic alternatives for dc coalitions) *Let $G \in 2^{Agt*}$. Then, $\mathcal{R}_G^{dc}$ is the binary relation on the set $\mathbf{S}$ such that, for all $B = (B_1, \ldots, B_n, V), B' = (B'_1, \ldots, B'_n, V') \in \mathbf{S}$:*

$$B\mathcal{R}_G^{dc}B' \text{ if and only if } \forall\alpha \in Comb_G(B) : B' \models \alpha.$$

[8] For more details about the differences between belief base combination and belief base merging, see [16].

Accessibility relations of type $\mathcal{R}_G^{dc}$ can be used to define a variant of the language $\mathcal{L}_1(Atm, Agt)$, denoted by $\mathcal{L}_3(Atm, Agt)$, for doxastically consistent coalitions. It is defined by the following grammar:

$$\varphi \quad ::= \quad \alpha \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_G^{dc}\varphi,$$

where α ranges over $\mathcal{L}_0(Atm, Agt)$ and G ranges over 2^{Agt*}.

The operator $\Box_G^{dc}\varphi$ has the following semantics:

$$(B, Cxt) \models \Box_G^{dc}\varphi \iff \forall B' \in Cxt : \text{if } B\mathcal{R}_G^{dc}B' \text{ then } (B', Cxt) \models \varphi.$$

As the following proposition highlights, in a complete model, a doxastically consistent coalition cannot implicitly believe a contradiction.

Proposition 10 *Let $G \in 2^{Agt*}$ and $B \in \mathbf{S}$. Then,*

$$(B, \mathbf{S}) \models \Diamond_G^{dc}\top,$$

where $\Diamond_G^{dc}\varphi \overset{\text{def}}{=} \neg\Box_G^{dc}\neg\varphi$.

PROOF. For every $X \in MCS_G(B)$, we have $||X||_{\mathbf{S}} \neq \emptyset$. Thus, by definition of $Comb_G(B)$, we have $||Comb_G(B)||_{\mathbf{S}} \neq \emptyset$. It follows that $\mathcal{R}_G^{dc}(B) \neq \emptyset$. Hence, $(B, \mathbf{S}) \models \Diamond_G^{dc}\top$. ■

As the following theorem indicates, model checking for language $\mathcal{L}_3$ has the same upper bound as model checking for $\mathcal{L}_1$ and $\mathcal{L}_2$.

Theorem 3 *Model checking and belief correct model checking for $\mathcal{L}_3$ are in PSPACE.*

SKETCH OF PROOF. To satisfy the hypothesis of Theorem 1, we only need to show that $Comb_G(B)$ can be computed in polynomial space in B. To compute $Comb_G(B)$, we consider all subsets $X \subseteq Pool_G(B)$ by lexicographic order. For each such X, we check whether it is consistent and any superset of it is not. Checking propositional consistency of a subset of $\mathcal{L}_0$-formulas is in NP, thus in PSPACE. So $Comb_G(B)$ can be computed in polynomial space. ■

In order to deal with potentially introspective coalitions, the notions of MCS and belief base combination need to be redefined, in order to warrant belief consistency. We denote by $MCS_G^{pi}(B)$ the set of maximally consistent subsets of a potentially introspective (pi) coalition G's belief base.

Definition 15 *Let $G \in 2^{Agt*}$ and $B \in \mathbf{S}$. Then, $X \in MCS_G^{pi}(B)$ if and only if:*

- *$X \subseteq Pool_G(B)$,*
- *$||X||_{(B,\mathbf{S})}^{\sim_G} \neq \emptyset$, and*
- *there is no $X' \subseteq Pool_G(B)$ s. th. $X \subset X'$ and $||X'||_{(B,\mathbf{S})}^{\sim_G} \neq \emptyset$,*

where, for every $Y \subseteq \mathcal{L}_0$,

$$||Y||_{(B,\mathbf{S})}^{\sim_G} = ||Y||_{\mathbf{S}} \cap \{B' \in \mathbf{S} : Pool_G(B) = Pool_G(B')\}.$$

Belief base combination for pi coalitions goes as follows.

Definition 16 (Combining for pi coalitions) *Let $B = (B_1, \ldots, B_n, V) \in \mathbf{S}$. Then,*

$$Comb_G^{pi}(B) = \bigcap_{X \in MCS_G^{pi}(B)} X.$$

$Comb_G^{pi}(B)$ is the result of combining the individual belief bases in B of the pi coalition G's members.

The following definition introduces the notion of doxastic accessibility relation for coalitions which are, at the same time, potentially introspective (pi) and doxastically consistent (dc).

Definition 17 (Doxastic alternatives for pi and dc coalitions) *Let $G \in 2^{Agt*}$. Then, $\mathcal{R}_G^{pi,dc}$ is the binary relation on the set $\mathbf{S}$ such that, for all $B, B' \in \mathbf{S}$, $B\mathcal{R}_G^{pi,dc}B'$ if and only if*

(i) $\forall\alpha \in Comb_G^{pi}(B) : B' \models \alpha$, and
(ii) $Pool_G(B) = Pool_G(B')$.

The following proposition is a variant of Proposition 7 for dc coalitions and is proved in a similar way.

Proposition 11 *Let $G \in 2^{Agt*}$. Then, the relation $\mathcal{R}_G^{pi,dc}$ is transitive and Euclidean.*

Accessibility relations of type $\mathcal{R}_G^{pi,dc}$ can be used to define a fourth variant of the language $\mathcal{L}_1(Atm, Agt)$, denoted by $\mathcal{L}_4(Atm, Agt)$, for potentially introspective, doxastically consistent coalitions. It is defined by the following grammar:

$$\varphi \quad ::= \quad \alpha \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \Box_G^{pi,dc}\varphi,$$

where α ranges over $\mathcal{L}_0(Atm, Agt)$ and G ranges over 2^{Agt*}. The operator $\Box_G^{pi,dc}\varphi$ has the following semantic interpretation:

$$(B, Cxt) \models \Box_G^{pi,dc}\varphi \iff \forall B' \in Cxt : \text{if } B\mathcal{R}_G^{pi,dc}B' \text{ then } (B', Cxt) \models \varphi.$$

The following proposition is the counterpart of Proposition 10 for pi coalitions: in a complete model, a potentially introspective, doxastically consistent coalition cannot implicitly believe a contradiction.

Proposition 12 *Let $G \in 2^{Agt*}$ and $B \in \mathbf{S}$. Then,*

$$(B, \mathbf{S}) \models \Diamond_G^{pi,dc}\top,$$

where $\Diamond_G^{pi,dc}\varphi \overset{\text{def}}{=} \neg\Box_G^{pi,dc}\neg\varphi$.

PROOF. For every $X \in MCS_G^{pi}(B)$, we have $||X||_{(B,\mathbf{S})}^{\sim_G} \neq \emptyset$. The latter means that there exists $B' \in \mathbf{S}$ such that $Pool_G(B) = Pool_G(B')$ and $(B', \mathbf{S}) \models \alpha$ for all $\alpha \in X$. Thus, by definition of $Comb_G^{pi}(B)$, there exists $B' \in \mathbf{S}$ such that $Pool_G(B) = Pool_G(B')$ and $(B', \mathbf{S}) \models \alpha$ for all $\alpha \in Comb_G^{pi}(B)$. It follows that $\mathcal{R}_G^{pi,dc}(B) \neq \emptyset$. Hence, $(B, \mathbf{S}) \models \Diamond_G^{pi,dc}\top$. ■

We conclude this section with PSPACE-membership of model checking for the language $\mathcal{L}_4$.

Theorem 4 *The model checking and belief correct model checking for $\mathcal{L}_4$ are in PSPACE.*

SKETCH OF PROOF. Again, to satisfy the hypothesis of Theorem 1, we only need to show that $Comb_G^{pi}(B)$ can be computed in polynomial space in B. In the spirit of the proof of Theorem 3, it is sufficient to show that checking $X \in MCS_G^{pi}(B)$ (Definition 15) can be turned into an algorithm in polynomial space. Let us just explain how to check that $||X||_{(B,\mathbf{S})}^{\sim_G} \neq \emptyset$. For that, we enumerate in polynomial space all B' with $Pool_G(B) = Pool_G(B')$ and accepts if formulas in X are true in B' (that can be checked in polynomial time since formulas in X are in $\mathcal{L}_0$). ■

8 Conclusion

Let's take stock. We have presented a new semantics for epistemic logic exploiting the concept of belief base and shown that it offers a natural framework for modelling the notion of distributed belief and, more generally, belief aggregation in a multi-agent setting. Unlike the belief merging approach that only deals with aggregation of propositional beliefs, our approach takes aggregation of higher-order beliefs into consideration. We have studied four notions of distributed belief, one based on belief pooling (i.e., $\Box_G$), its variant for potentially introspective coalitions (i.e., $\Box_G^{pi}$), another one based on belief combination which guarantees collective belief consistency (i.e., $\Box_G^{dc}$), and its variant for potentially introspective coalitions (i.e., $\Box_G^{pi,dc}$).

The logical framework we presented is static, as the agents' belief bases do not change. Future work will be devoted to bringing into our framework different types of belief change operations including belief base expansion, contraction and revision. The extended framework will allow us to study the interplay between belief change and belief aggregation. For instance, following [28, 30], we plan to investigate whether it should be possible for the members of a coalition to come to individually believe that a certain fact φ is true, through information exchange, when they distributively believe that φ.

ACKNOWLEDGEMENTS

Support from the ANR-3IA Artificial and Natural Intelligence Toulouse Institute and from the ANR project CoPains ("Cognitive Planning in Persuasive Multimodal Communication", grant number ANR-18-CE33-0012) is gratefully acknowledged.

References

[1] T. Ågotnes and Y. N. Wáng, 'Resolving distributed knowledge', *Artificial Intelligence*, **252**, 1–21, (2017).
[2] C. Alchourrón, P. Gärdenfors, and D. Makinson, 'On the logic of theory change: Partial meet contraction and revision functions', *The Journal of Symbolic Logic*, **50**, 510–530, (1985).
[3] S. E. Asch, 'Studies of independence and conformity: a minority of one against a unanimous majority', *Psychological Monographs*, **70**(9), 255–273, (1956).
[4] C. Baral, S. Kraus, and J. Minker, 'Combining multiple knowledge bases', *IEEE Transactions Knowledge Data Engineering*, **3**(2), 208–220, (1991).
[5] C. Baral, S. Kraus, J. Minker, and V. S. Subrahmanian, 'Combining knowledge bases consisting of first-order analysis', *Computational Intelligence*, **8**, 45–71, (1992).
[6] T. Bolander and M. B. Andersen, 'Epistemic planning for single- and multi-agent systems', *Journal of Applied Non-Classical Logics*, **21**(1), 656–680, (2011).
[7] M. Burrows, M. Abadi, and R. Needham, 'A logic of authentication', *ACM Transactions on Computer Systems*, **8**(1), 18–36, (1990).
[8] D. C. Dennett, *The Intentional Stance*, MIT Press, Cambridge, Massachusetts, 1987.
[9] R. Fagin, J. Halpern, Y. Moses, and M. Vardi, *Reasoning about Knowledge*, MIT Press, Cambridge, 1995.
[10] L. Festinger, 'Informal social communication', *Psychological Review*, **57**(5), 271–282, (1950).
[11] T. French, W. van der Hoek, P. Iliev, and B. P. Kooi, 'On the succinctness of some modal logics', *Artificial Intelligence*, **197**, 56–85, (2013).
[12] J. Y. Halpern and Y. Moses, 'A guide to completeness and complexity for modal logics of knowledge and belief', *Artificial Intelligence*, **54**(2), 319–379, (1992).
[13] J. Y. Halpern and R. Pass, 'A knowledge-based analysis of the blockchain protocol', in *Proceedings of the Sixteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2017)*, pp. 324–335. EPTCS, (2017).
[14] A. Herzig and F. Maffre, 'How to share knowledge by gossiping', *AI Communications*, **30**(1), 1–17, (2017).
[15] J. Hintikka, *Knowledge and Belief*, Cornell University Press, New York, 1962.
[16] S. Konieczny, 'On the difference between merging knowledge bases and combining them', in *Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning (KR 2000)*, eds., A. G. Cohn, F. Giunchiglia, and B. Selman, pp. 135–144. Morgan Kaufmann, (2000).
[17] S. Konieczny and R. Pino Pérez, 'Merging information under constraints: A logical framework', *Journal of Logic and Computation*, **12**(5), 773–808, (2002).
[18] K. Konolige, *A deduction model of belief*, Morgan Kaufmann Publishers, Los Altos, 1986.
[19] L. Lismont and P. Mongin, 'On the logic of common belief and common knowledge', *Theory and Decision*, **37**, 75–106, (1994).
[20] E. Lorini, 'In praise of belief bases: Doing epistemic logic without possible worlds', in *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18)*, pp. 1915–1922. AAAI Press, (2018).
[21] E. Lorini, 'Exploiting belief bases for building rich epistemic structures', in *Proceedings of the Seventeenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2019)*, ed., L. S. Moss, volume 297 of *EPTCS*, pp. 332–353, (2019).
[22] E. Lorini, 'Rethinking epistemic logic with belief bases', *Artificial Intelligence*, **282**, (2020).
[23] E. Lorini, D. Longin, B. Gaudou, and A. Herzig, 'The logic of acceptance: Grounding institutions on agents' attitudes', *Journal of Logic and Computation*, **19**(6), 901–940, (2009).
[24] E. Lorini and F. Romero, 'Decision procedures for epistemic logic exploiting belief bases', in *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS 2019)*, pp. 944–952. IFAAMAS, (2019).
[25] B. Marinkovic, P. Glavan, Z. Ognjanovic, and T. Studer, 'A temporal epistemic logic with a non-rigid set of agents for analyzing the blockchain protocol', *Journal of Logic and Computation*, **29**(5), 803–830, (2019).
[26] J. J. Meyer and W. van der Hoek, *Epistemic logic for AI and computer science*, Cambridge University Press, 1995.
[27] G. Pigozzi, 'Belief merging and judgment aggregation', *Stanford Encyclopedia of Philosophy*, (2015).
[28] F. Roelofsen, 'Distributed knowledge', *Journal of Applied Non-Classical Logics*, **17**(2), 255–273, (2007).
[29] R. Stalnaker, 'Common ground', *Linguistics and Philosophy*, **25(5-6)**, 701–721, (2002).
[30] W. van Der Hoek, B. van Linder, and J.-J. Meyer, 'Group knowledge is not always distributed (neither is it always implicit)', *Mathematical social sciences*, **38**(2), 215–240, (1999).
[31] P. Vanderschraaf and G. Sillari, 'Common knowledge', *Stanford Encyclopedia of Philosophy*, (2011).
[32] Y. N. Wáng and T. Ågotnes, 'Public announcement logic with distributed knowledge: expressivity, completeness and complexity', *Synthese*, **190**(1), 135–162, (2013).
[33] R. H. Willis, 'Two dimensions of conformity-nonconformity', *Sociometry*, **26**, 499–513, (1963).