Towards Epistemic-Doxastic Planning with Observation and Revision

Thorsten Engesser[1], Andreas Herzig[2], Elise Perrotin[3]

[1]IRIT, Toulouse, France
[2]IRIT, CNRS, Toulouse, France
[3]CRIL, CNRS, Lens, France
thorsten.engesser@irit.fr, andreas.herzig@irit.fr, perrotin@cril.fr

Abstract

Epistemic planning is useful in situations where multiple agents have different knowledge and beliefs about the world, such as in robot-human interaction. One aspect that has been largely neglected in the literature is planning with observations in the presence of false beliefs. This is a particularly challenging problem because it requires belief revision. We introduce a simple specification language for reasoning about actions with knowledge and belief. We demonstrate our approach on well-known false-belief tasks such as the Sally-Anne Task and compare it to other action languages. Our logic leads to an epistemic planning formalism that is expressive enough to model second-order false-belief tasks, yet has the same computational complexity as classical planning.

1 Introduction

The ability to perform theory of mind reasoning is a valuable skill for autonomous agents. For example, planning from the perspective of other agents can facilitate coordination between agents (Engesser et al. 2017; Bolander et al. 2018). It is also fundamental in robot-human interaction scenarios where the robot has to account for (and potentially adjust) a human's false beliefs (Dissing and Bolander 2020).

Most epistemic planning approaches in the literature use either knowledge or belief, but not both. An exception is the work of Andersen, Bolander, and Jensen (2015), which however does not focus on false beliefs but rather on the plausibility of action outcomes. We argue that the distinction between knowledge and belief is relevant to the applications mentioned above: basing decisions on knowledge provides more guarantees than basing them on belief.

Another problem shared by many existing epistemic planning formalisms is their complexity. In particular, planning based on Dynamic Epistemic Logic (DEL) is undecidable (Bolander and Andersen 2011; Aucher and Bolander 2013). Some useful PSPACE-complete fragments of DEL-based planning were identified (Kominis and Geffner 2015; Engesser and Miller 2020; Bolander et al. 2020), but they typically disallow actions increasing the agents' uncertainty. Among other things, this makes it impossible to model most false-belief tasks from the literature.

Copyright © 2024, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Static epistemic reasoning is actually already difficult: deciding satisfiability for epistemic formulas is PSPACE-hard if there is more than one agent, and EXPTIME-complete if the language contains the common knowledge operator (Fagin et al. 1995). Several authors have proposed lightweight fragments of the epistemic language, most prominently in terms of epistemic literals and proper epistemic knowledge bases (Lakemeyer and Lespérance 2012). These fragments forbid conjunctions and disjunctions in the scope of the epistemic operators. Formulas are therefore boolean combinations of what may be called *epistemic literals*: sequences of 'knowing-that' operators $\mathtt{K}_i$ and $\widehat{\mathtt{K}}_i$ followed by a propositional variable or its negation. Static epistemic reasoning then typically reduces to propositional reasoning, and epistemic planning to classical planning (Muise et al. 2022). The reduction considers epistemic literals as arbitrary propositional literals without any structure. This requires an axiomatisation of their interaction. For example, the conjunction $\mathtt{K}_i p \wedge \mathtt{K}_i \neg p$ is unsatisfiable in epistemic logic but is propositionally satisfiable: one has to add its negation as an axiom in order to obtain propositional unsatisfiability.

An alternative route was proposed by Herzig, Lorini, and Maffre (2018) and Cooper et al. (2021) in terms of the 'knowing whether' (or 'knowledge about') operator $\mathtt{KA}_i$. There, an epistemic atom is a sequence of $\mathtt{KA}_i$ operators followed by a propositional variable. An advantage over 'knowledge that' literals is that less interactions have to be axiomatised. If the epistemic logic is S5, then the epistemic atoms $\mathtt{KA}_i\mathtt{KA}_i p$ are tautologous; more generally, this is the case for *epistemic atoms with repetitions* which have the form $\cdots \mathtt{KA}_i\mathtt{KA}_i \cdots p$. Luckily, this is the only axiom to be taken into account: if we restrict all atoms to be *repetition-free*, then all atoms are logically independent and epistemic reasoning reduces to propositional reasoning.

We here generalise the 'knowledge about' approach by adding belief. Our epistemic-doxastic atoms consist of sequences of modal operators $\mathtt{TBA}_i$ and $\mathtt{MBA}_i$, followed by a propositional variable, where $\mathtt{TBA}_i\varphi$ reads "i has a true belief about φ" and $\mathtt{MBA}_i\varphi$ reads "i has a mere belief about φ". We prove that reasoning with boolean combinations of repetition-free epistemic-doxastic atoms can be directly done in propositional logic. In that lightweight fragment we define laws for ontic and epistemic actions. The latter are of two kinds: starting and ceasing to observe an atom, where we identify observation of an atom with knowledge about that atom. This allows us to track the agents' knowledge and beliefs in scenarios involving higher-order false beliefs. Our logic is the basis for a useful multiagent planning formalism that has the same complexity as classical planning and that is suitable for modelling, among other things, second-order false-belief tasks.

The paper is organised as follows. After recalling epistemic-doxastic logic (Section 2), we define our lightweight fragment (Section 3) and make explicit the hypotheses about observations and inertia of knowledge and belief (Section 4). We then introduce ontic and epistemic actions (Sections 5 and 6) and illustrate them on a second-order false-belief task (Section 7). We define epistemic-doxastic planning (Section 8), and finally we compare our approach to action languages from related work and discuss its limitations (Section 9).

2 Background

Throughout the paper $\mathbb{P}$ is a countable set of propositional variables, with typical elements $p, q, \ldots$; and $\mathbb{A}$ is a countable set of agents, with typical elements $i, j, \ldots$

In standard presentations, the vocabulary from which formulas are built is identical to the set $\mathbb{P}$. We here consider more generally that the vocabulary is a countable set Voc that may have some structure. Then a *propositional formula on* Voc is a boolean combination of atoms from Voc. The propositional language on Voc, noted $\mathcal{L}_{\mathsf{bool}}(\mathsf{Voc})$, is the set of all boolean combinations of elements of Voc.

Propositional valuations on Voc, alias *states*, are subsets of Voc. Propositional satisfaction of a formula $\varphi \in \mathcal{L}_{\mathsf{bool}}(\mathsf{Voc})$ in a valuation $V \subseteq \mathsf{Voc}$ is denoted by $V \models \varphi$ and defined as usual. A formula $\varphi \in \mathcal{L}_{\mathsf{bool}}(\mathsf{Voc})$ is *propositionally valid* if $V \models \varphi$ for every $V \subseteq \mathsf{Voc}$. It is *propositionally satisfiable* if $\neg\varphi$ is not propositionally valid.

Epistemic-Doxastic Logic S5-EDL: Kripke Models

We suppose that knowledge comes from observations, and that these observations are reliable. This justifies the choice of the logic S5 and in particular that of the negative introspection axiom $\neg\mathtt{K}_i\varphi \to \mathtt{K}_i\neg\mathtt{K}_i\varphi$ (which can be criticised otherwise, cf. Voorbraak 1993, Aucher 2015).

An *S5-EDL model* is a quadruple $M = \langle W, \{K_i\}_{i\in\mathbb{A}}, \{B_i\}_{i\in\mathbb{A}}, val\rangle$ where W is a non-empty set of possible worlds; every K_i and B_i is a binary relation on W; and $val : W \to 2^{\mathbb{P}}$ associates propositional valuations on $\mathbb{P}$ to possible worlds. The accessibility relations K_i and B_i must satisfy the following constraints:

- Each K_i is an equivalence relation;
- Each B_i is serial, transitive, and euclidean;
- $B_i \subseteq K_i$, for every $i \in \mathbb{A}$;
- $K_i \circ B_i \subseteq B_i$, for every $i \in \mathbb{A}$.

Knowing-That and Believing-That

The standard language of epistemic-doxastic logic is defined by the following grammar:

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathtt{K}_i\varphi \mid \mathtt{B}_i\varphi,$$

where p ranges over $\mathbb{P}$ and i ranges over $\mathbb{A}$. The formula $\mathtt{K}_i\varphi$ reads "i knows that φ" and $\mathtt{B}_i\varphi$ reads "i believes that φ". The truth conditions are:

$$\begin{array}{lll} M, w \Vdash p & \text{if} & p \in val(w), \text{for } p \in \mathbb{P}; \\ M, w \Vdash \mathtt{K}_i\varphi & \text{if} & \text{for every } w' \in K_i(w), M, w' \Vdash \varphi; \\ M, w \Vdash \mathtt{B}_i\varphi & \text{if} & \text{for every } w' \in B_i(w), M, w' \Vdash \varphi; \end{array}$$

and as usual for the boolean connectives, where $K_i(w) = \{w' : \langle w, w'\rangle \in K_i\}$ and likewise for $B_i(w)$. Validity and satisfiability are defined in the standard way. The logic of the $\mathtt{K}_i$ operators is S5; the logic of the $\mathtt{B}_i$ is KD45; the principles $\mathtt{B}_i\varphi \to \mathtt{K}_i\mathtt{B}_i\varphi$ and $\mathtt{K}_i\varphi \to \mathtt{B}_i\varphi$ complete the axiomatisation. Moreover, the principles $\neg\mathtt{B}_i\varphi \to \mathtt{K}_i\neg\mathtt{B}_i\varphi$ and $\mathtt{B}_i\mathtt{K}_i\varphi \to \mathtt{K}_i\varphi$ are valid (Voorbraak 1993; Aucher 2015). Just like negative introspection for $\mathtt{K}_i$, the last validity is not acceptable in general, but is so for knowledge based on reliable observation.

True Belief and Mere Belief

We now give an alternative presentation of S5-EDL that follows Herzig and Perrotin (2021).

Epistemic-doxastic formulas are defined by the grammar

$$\varphi ::= p \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathtt{TBA}_i\varphi \mid \mathtt{MBA}_i\varphi,$$

where p ranges over $\mathbb{P}$ and i over $\mathbb{A}$. The formula $\mathtt{TBA}_i\varphi$ reads "i has a true belief about φ" and $\mathtt{MBA}_i\varphi$ reads "i has a mere belief about φ". The *epistemic-doxastic language* is noted $\mathcal{L}_{\mathrm{EDL}}(\mathbb{P})$.

We say that a set of worlds $S \subseteq W$ of a model M *agrees* on φ if either $M, w \Vdash \varphi$ for every $w \in S$, or $M, w \nVdash \varphi$ for every $w \in S$; otherwise we say that S *disagrees* on φ. Then:

$$\begin{array}{lll} M, w \Vdash \mathtt{TBA}_i\varphi & \text{if} & B_i(w) \cup \{w\} \text{ agrees on } \varphi; \\ M, w \Vdash \mathtt{MBA}_i\varphi & \text{if} & B_i(w) \text{ agrees on } \varphi, \text{ and} \\ & & K_i(w) \text{ disagrees on } \varphi. \end{array}$$

Validity and satisfiability are defined as before.

The language $\mathcal{L}_{\mathrm{EDL}}(\mathbb{P})$ has the same expressivity as the standard epistemic-doxastic language: in one direction, we define $\mathtt{TBA}_i\varphi$ as $(\varphi \wedge \mathtt{B}_i\varphi) \vee (\neg\varphi \wedge \mathtt{B}_i\neg\varphi)$ and $\mathtt{MBA}_i\varphi$ as $(\mathtt{B}_i\varphi \wedge \neg\mathtt{K}_i\varphi) \vee (\mathtt{B}_i\neg\varphi \wedge \neg\mathtt{K}_i\neg\varphi)$; in the other direction, we define $\mathtt{K}_i\varphi$ as $\mathtt{TBA}_i\varphi \wedge \neg\mathtt{MBA}_i\varphi$ and $\mathtt{B}_i\varphi$ as $(\varphi \wedge \mathtt{TBA}_i\varphi) \vee (\neg\varphi \wedge \neg\mathtt{TBA}_i\varphi \wedge \mathtt{MBA}_i\varphi)$.

The logical combinations of $\mathtt{TBA}_i\varphi$ and $\mathtt{MBA}_i\varphi$ account for all possible epistemic-doxastic attitudes of i w.r.t. φ:

$$\begin{array}{ll} \mathtt{OBS}_i\varphi = \mathtt{TBA}_i\varphi \wedge \neg\mathtt{MBA}_i\varphi & \text{(observation)} \\ \mathtt{LBA}_i\varphi = \mathtt{TBA}_i\varphi \wedge \mathtt{MBA}_i\varphi & \text{(lucky belief)} \\ \mathtt{FBA}_i\varphi = \neg\mathtt{TBA}_i\varphi \wedge \mathtt{MBA}_i\varphi & \text{(false belief)} \\ \mathtt{NBA}_i\varphi = \neg\mathtt{TBA}_i\varphi \wedge \neg\mathtt{MBA}_i\varphi & \text{(no belief)} \end{array}$$

An agent observes φ if and only if she has knowledge about φ. Hence $\mathtt{OBS}_i\varphi$ is nothing but $\mathtt{KA}_i\varphi$; we however prefer the present notation because it matches our hypothesis that knowledge comes from observation.[1]

[1] Lucky beliefs are debated in the epistemology literature: can we give conditions under which knowledge can be separated from 'epistemic luck', i.e., luckily believing a proposition (Ichikawa and Steup 2018)? We here suppose that the distinction can be made.

$$\begin{tikzcd} \boxed{\begin{array}{c} p,\\ \mathtt{TBA_S}p, \neg\mathtt{MBA_S}p,\\ \mathtt{TBA_A}p, \neg\mathtt{MBA_A}p \end{array}} \arrow[r, "\text{Sally stops observing}"] & \boxed{\begin{array}{c} p,\\ \mathtt{TBA_S}p, \underline{\mathtt{MBA_S}p},\\ \mathtt{TBA_A}p, \neg\mathtt{MBA_A}p \end{array}} \arrow[r, "\text{Anne moves the marble}"] & \boxed{\begin{array}{c} \underline{\neg p},\\ \underline{\neg\mathtt{TBA_S}p}, \mathtt{MBA_S}p,\\ \mathtt{TBA_A}p, \neg\mathtt{MBA_A}p \end{array}} \end{tikzcd}$$

$\models p \land \mathtt{OBS_S}p \land \mathtt{OBS_A}p$ $\models p \land \mathtt{LBA_S}p \land \mathtt{OBS_A}p$ $\models \neg p \land \mathtt{FBA_S}p \land \mathtt{OBS_A}p$

Figure 1: The Sally-Anne Task. Atoms that have changed from the previous state are underlined.

3 Repetition-Free Epistemic-Doxastic Atoms

We now define a lightweight fragment of the language $\mathcal{L}_{\text{EDL}}(\mathbb{P})$. It consists in boolean combinations of modal atoms: sequences of modal operators $\mathtt{TBA}_i$ and $\mathtt{MBA}_i$ followed by a propositional variable.

Definition of the Fragment

If i is an agent, an *i-formula* is a formula of $\mathcal{L}_{\text{EDL}}(\mathbb{P})$ of the form either $\mathtt{TBA}_i\varphi$ or $\mathtt{MBA}_i\varphi$. The set of repetition-free epistemic-doxastic atoms *REDA* is the smallest set such that:

- $\mathbb{P} \subseteq \mathit{REDA}$;
- If $i \in \mathbb{A}$, $\alpha \in \mathit{REDA}$ and α is not an i-formula then $\mathtt{TBA}_i\alpha \in \mathit{REDA}$ and $\mathtt{MBA}_i\alpha \in \mathit{REDA}$.

For example, $\mathtt{TBA}_i\mathtt{MBA}_j p$ is in *REDA* iff $i \neq j$. The propositional language on *REDA*, $\mathcal{L}_{\mathsf{bool}}(\mathit{REDA})$, is the set of boolean combinations of *REDA* atoms. It can also be viewed as a fragment of the epistemic-doxastic language on $\mathbb{P}$: we have that $\mathcal{L}_{\mathsf{bool}}(\mathit{REDA}) \subseteq \mathcal{L}_{\text{EDL}}(\mathbb{P})$. The set of atoms of modal depth at least one is noted $\mathit{REDA}^{\geq 1}$; $\mathit{REDA}^{\leq 1}$ is that of depth at most one; and $\mathit{REDA}^{\leq 2}$ that of depth at most two.

Example: the Sally-Anne Task

False-belief tasks are well-known experiments from cognitive psychology to determine the ability of people—usually children—to use *theory-of-mind* reasoning. The paradigmatic *Sally-Anne Task* goes as follows:

> Two children, Sally and Anne, are in a room together. Sally has a marble, which she puts into a basket. Then she leaves the room to go out for a walk. While Sally is away, Anne removes the marble from the basket and puts it into a box (which is also in the room). Finally, Sally comes back into the room. Will Sally search for her marble in the basket or in the box?

The answer requires reasoning from Sally's perspective. It is well-known that children under the age of four as well as older children with autism spectrum disorder have difficulties performing this kind of reasoning (Wimmer and Perner 1983; Baron-Cohen, Leslie, and Frith 1985).

Figure 1 depicts the evolution of facts, knowledge, and beliefs in terms of *REDA* valuations. We use a single propositional variable p: if it is true then the marble is in the basket; if it is false then the marble is in the box. Furthermore, we restrict our attention to first-order beliefs about p.

Initially, the marble is in the basket, and both agents observe this: they have true beliefs about p that are not mere beliefs. After Sally has left she can no longer observe the marble. This means that her belief about the marble becomes a mere belief. Since that belief is still true, Sally now has a lucky belief about the position of the marble. Finally, when Anne moves the marble, the value of the proposition p changes. Since Sally does not observe p (which is clear from the fact that $\mathtt{MBA_S}p$ is true), her belief about the position of the marble should not change. Thus, her lucky belief must become a false belief. When Sally returns (assuming that returning does not further change Sally's observations), she falsely believes that the marble is still in the basket.

Equivalence with Propositional Satisfiability

In the rest of the section we prove that for the fragment $\mathcal{L}_{\mathsf{bool}}(\mathit{REDA})$ of $\mathcal{L}_{\text{EDL}}(\mathbb{P})$, S5-EDL satisfiability is equivalent to propositional satisfiability in a valuation $V \subseteq \mathit{REDA}$.

Theorem 1 *For every $\varphi \in \mathcal{L}_{\mathsf{bool}}(\mathit{REDA})$, φ is S5-EDL satisfiable iff φ is propositionally satisfiable.*

PROOF. The left-to-right direction is straightforward: given a pointed model (M, w), we extract a valuation $V_w = \{\alpha \in \mathit{REDA} : M, w \Vdash \alpha\}$ and show by induction that any formula $\varphi \in \mathcal{L}_{\mathsf{bool}}(\mathit{REDA})$ is propositionally true in V_w if and only if it is true in (M, w).

The right-to-left direction relies on the construction of a canonical model $M^c = \langle 2^{\mathit{REDA}}, \{K_i^c\}_{i \in \mathbb{A}}, \{B_i^c\}_{i \in \mathbb{A}}, \mathit{val}^c \rangle$ such that for any $V \subseteq \mathit{REDA}$ and any formula $\varphi \in \mathcal{L}_{\mathsf{bool}}(\mathit{REDA})$ we have that $V \models \varphi$ iff $M^c, V \Vdash \varphi$. The delicate part is defining the relations K_i^c and B_i^c (val^c is defined naturally as $\mathit{val}^c(V) = V \cap \mathbb{P}$). If V and U are subsets of *REDA*, let $V K_i^c U$ iff for all $\alpha \in \mathit{REDA}$:

- $\mathtt{MBA}_i\alpha \in V$ iff $\mathtt{MBA}_i\alpha \in U$;
- If $\mathtt{MBA}_i\alpha \notin V$ then $\mathtt{TBA}_i\alpha \in V$ iff $\mathtt{TBA}_i\alpha \in U$;
- If $\mathtt{TBA}_i\alpha \in V \cap U$ then V and U agree on α;
- If $\mathtt{MBA}_i\alpha \in V$ and $\mathtt{TBA}_i\alpha \notin V \cup U$ then V, U agree on α;
- If $\mathtt{TBA}_i\alpha \in (V \setminus U) \cup (U \setminus V)$ then V, U disagree on α.

Intuitively, the first two conditions ensure that agent i has knowledge and belief about the same atoms in V and U, while the latter three ensure that those knowledge and beliefs have the same truth value for those atoms (e.g. if an agent believes p in V then she also believes p, and not $\neg p$, in U).

If V and U are subsets of *REDA*, we further say that $V B_i^c U$ if $V K_i^c U$ and for all $\alpha \in \mathit{REDA}$, if $\mathtt{MBA}_i\alpha \in V$ then $\mathtt{TBA}_i\alpha \in U$. That is, agents believe they are in a world in which all of their beliefs are correct. From there, it is relatively straightforward to show that M^c is an S5-EDL model, that is, that K_i^c and B_i^c have the right properties. ■

4 Adding Dynamics

We now add actions to the picture. We do so in a principled way: we first spell out some hypotheses and then define the general format of action descriptions and their semantics.

H1: All knowledge comes from observation. We suppose that agents' knowledge comes from observations. As already argued, this justifies the choice of S5 as the logic of knowledge. It follows that if an agent i has knowledge about α then this not only means that i observes the truth value of α (and therefore either knows that α or that $\neg\alpha$), but also that i observes any change in the truth value of α.

In terms of our logical language of true and mere beliefs, "i observes α" is expressed by $\mathtt{OBS}_i\alpha = \mathtt{TBA}_i\alpha \wedge \neg\mathtt{MBA}_i\alpha$. If the truth value of α changes then i's observation means that $\mathtt{TBA}_i\alpha$ cannot become false[2] and that $\mathtt{MBA}_i\alpha$ cannot become true; that is, $\mathtt{OBS}_i\alpha$ remains true.

H2: Lack of observation implies inertia of beliefs. When an agent does not observe α, she either holds a (possibly false) mere belief about α or has no belief about α. In both cases, we suppose that a change of α does not affect the status of her belief about α. Thus, an agent cannot acquire a new belief without observation.[3] We also exclude that agents lose beliefs.[4] It follows that an agent continues to believe the last truth value of α that she has observed, and so until a new observation tells her otherwise. It also follows that when an agent has never observed α, she cannot but have no belief about α.

If i does not observe α then there are three possibilities: $\mathtt{NBA}_i\alpha$ (no belief about α), $\mathtt{LBA}_i\alpha$ (lucky belief about α), and $\mathtt{FBA}_i\alpha$ (false belief about α). If an action flips α then:

- If $\mathtt{NBA}_i\alpha$ was true, then $\mathtt{NBA}_i\alpha$ remains true.
- If $\mathtt{LBA}_i\alpha$ was true, then $\mathtt{FBA}_i\alpha$ will be true afterwards.
- If $\mathtt{FBA}_i\alpha$ was true, then $\mathtt{LBA}_i\alpha$ will be true afterwards.

So a mere belief about α that happens to be false becomes a lucky belief about α when α flips, and vice versa.

To sum it up, a change of α has the following effects in terms of true and mere beliefs: (a) $\mathtt{MBA}_i\alpha$ is stable: if it was true then it remains true, and if it was false then it remains false[5]; (b) if $\mathtt{MBA}_i\alpha$ was true then $\mathtt{TBA}_i\alpha$ is flipped.

H3: Actions are either ontic or epistemic. Ontic actions are actions whose primary effects are about the world, such as moving an object. These effects therefore concern subsets of the set of propositional variables $\mathbb{P}$. In contrast, epistemic actions have no effects on the physical world: all effects concern the agents' beliefs, and the atoms changed by an epistemic action are elements of $\mathit{REDA}^{\geq 1}$. The only action that is both ontic and epistemic is the 'do nothing' action.

[2]This supposes that no action can flip both α and $\mathtt{TBA}_i\alpha$. The actions that we are going to consider satisfy that constraint, as we make explicit in our hypothesis (H3).

[3]We therefore exclude the possibility of doxastic voluntarism, an issue that is much debated in philosophy (Chignell 2018).

[4]We are aware that this is a strong hypothesis because the strength of an agent's belief about something she doesn't observe typically decreases over time. In order to obtain a more realistic account one has to add more machinery and introduce things such as forgetting, deadlines for belief (belief that α turns into ignorance about α after some time), and, ultimately, degrees of belief.

[5]In the latter case, if $\mathtt{TBA}_i\alpha$ holds then i observes α, and stability already follows from (H1).

H4: Higher-order effects are limited to depth at most two. In order to simplify things, we restrict the kind of effects that we consider: we suppose that all atoms involved in action effects are of modal depth at most two. The atoms that are changed by an epistemic action are therefore elements of $\mathit{REDA}^{\leq 2}$. In consequence, we also restrict higher-order reasoning about knowledge and belief: description of the agents' knowledge and belief is supposed to be at most of depth two as well; and states are simply subsets of $\mathit{REDA}^{\leq 2}$.

Direct and Indirect Effects of an Action

Actions are described by sets of conditional effects of the form $\varphi \triangleright \alpha$, where φ is a formula from $\mathcal{L}_{\mathsf{bool}}(\mathit{REDA})$ and α is an element of $\mathit{REDA}^{\leq 2}$. The intuition is that if the condition φ is true then the truth value of α is *flipped* by the action. For example, the ontic action of removing the marble from the basket has the single conditional effect $p \triangleright p$. So if p is true, then the truth value of p gets flipped, that is, it becomes false. As this is the only conditional effect, the truth value of p remains false when p is false.[6]

The set of all conditional effects of action a is noted $\mathit{eff}(a)$. We partition $\mathit{eff}(a)$ into direct effects $\mathit{deff}(a)$ and indirect effects $\mathit{ieff}(a)$. The indirect effects are derived from the direct effects and are the action's effects on the agents' knowledge and beliefs. They are conditioned by the agents' observational status w.r.t. the direct effects. In particular, the following principle follows from our hypothesis H2:

$$\text{If } \varphi \triangleright \alpha \in \mathit{eff}(a) \text{ then } \varphi \wedge \mathtt{MBA}_i\alpha \triangleright \mathtt{TBA}_i\alpha \in \mathit{eff}(a). \quad \text{(M)}$$

In words: if, under circumstances described by φ, α flips and i has a mere belief about α, then $\mathtt{TBA}_i\alpha$ also flips. The only situation where (M) does not apply is when $\mathtt{TBA}_i\alpha$ is among the direct effects of a. In our framework this will only be the case for the action type of starting shared observation.

Principle (M) allows us to derive an effect of epistemic level $n+1$ from an effect at level n. There is a similar principle deriving an effect of epistemic level $n+2$ from an effect at level n. Due to our restriction to epistemic depth at most 2 it only applies to ontic actions. We therefore postpone the discussion to the next section, where we associate indirect effects to each type of action. For each of these we will argue that no other indirect effects need to be considered.

Semantics of an Action

Let $\mathit{eff}(a)$ be the effects of action a. The result of applying a to a state V is $a(V) =$

$$\{\alpha \in V \ : \text{ there is no } \varphi \triangleright \alpha \in \mathit{eff}(a) \text{ such that } V \models \varphi\} \cup$$
$$\{\alpha \notin V \ : \text{ there is a } \varphi \triangleright \alpha \in \mathit{eff}(a) \text{ such that } V \models \varphi\}.$$

We have $a(V) \subseteq \mathit{REDA}^{\leq 2}$ because (1) $V \subseteq \mathit{REDA}^{\leq 2}$ and (2) all action effects are restricted to $\mathit{REDA}^{\leq 2}$. With $V \subseteq \mathit{REDA}^{\leq 2}$, no atoms of depth greater than 2 need to be added by Principle (M), or any reasonable higher order principle.

[6]Any STRIPS action with add-list $P^+ \subseteq \mathbb{P}$ and delete-list $P^- \subseteq \mathbb{P}$ can be expressed in that format as an action a with conditional effects $\neg p \triangleright p$ for every $p \in P^+$ and $p \triangleright p$ for every $p \in P^-$. We could add action preconditions as usually done in the planning literature; we will do so in Section 8.

5 Ontic Actions

The direct effects of an ontic action a are described by a finite set of conditional effects of the form $\varphi \triangleright p$, for $p \in \mathbb{P}$. Let us discuss how indirect effects are computed from the direct effects. Consider the ontic action $\mathtt{flip}(p)$ of changing the truth value of the propositional variable p. That is, $\mathit{deff}(\mathtt{flip}(p)) = \{\top \triangleright p\}$. According to H1, the first-order indirect effects of $\mathtt{flip}(p)$ should be that all agents who observe p know that p's truth value got flipped. According to H2, all agents not observing p stick to their previous belief: if before the action they believed that p then after the action they still believe that p; and if they believed $\neg p$ then after the action they still believe $\neg p$. Hence the truth status of mere belief flips between lucky belief and false belief, that is,

$$\mathtt{MBA}_i p \triangleright \mathtt{TBA}_i p.$$

This is a consequence of Principle (M) of Section 4. This effect of $\mathtt{flip}(p)$ on all $\mathtt{TBA}_i p$ is its only first-order indirect effect: the truth value of all $\mathtt{MBA}_i p$ is stable because ontic actions do not modify the agents' observability, i.e., what the agents do and don't observe.

Let us turn to the second-order indirect effects. First of all, the application of Principle (M) allows us to derive a second-order indirect effect of $\mathtt{flip}(p)$ from the first-order indirect effect $\mathtt{MBA}_i p \triangleright \mathtt{TBA}_i p$, namely

$$\mathtt{MBA}_i p \wedge \mathtt{MBA}_j \mathtt{TBA}_i p \ \triangleright\ \mathtt{TBA}_j \mathtt{TBA}_i p.$$

In words: if i does not observe p but holds a belief about p then, as $\mathtt{flip}(p)$ also flips $\mathtt{TBA}_i p$, the truth value of $\mathtt{TBA}_j \mathtt{TBA}_i p$ must also be flipped as soon as j holds a mere belief about $\mathtt{TBA}_i p$.

The preceding second-order indirect effect occurs when $\mathtt{TBA}_i p$ is flipped while agent j, due to a mere belief about $\mathtt{TBA}_i p$, wrongly believes it is not going to change. There is also a symmetric case where $\mathtt{TBA}_i p$ is stable while j wrongly believes that it is going to be flipped. This requires a situation where j wrongly believes that $\mathtt{MBA}_i p$ is true; moreover, j must observe the change of p in order to mistakenly deduce that $\mathtt{TBA}_i p$ will be flipped. Hence $\mathtt{flip}(p)$ has the following second-order indirect effect:

$$\neg\mathtt{MBA}_i p \wedge \mathtt{FBA}_j \mathtt{MBA}_i p \wedge \mathtt{OBS}_j p \ \triangleright\ \mathtt{TBA}_j \mathtt{TBA}_i p.$$

These two second-order indirect effects of $\mathtt{flip}(p)$ on $\mathtt{TBA}_j \mathtt{TBA}_i p$ are the only ones: the truth values of all $\mathtt{TBA}_j \mathtt{MBA}_i p$, $\mathtt{MBA}_j \mathtt{TBA}_i p$ and $\mathtt{MBA}_j \mathtt{MBA}_i p$ are stable because ontic actions do not modify the agents' observability, and all agents know that and know that all agents know that.

Altogether, we obtain the following set of indirect conditional effects of an ontic action a:

$$\begin{aligned}
&\mathit{ieff}(a) = \\
&\{\varphi \wedge \mathtt{MBA}_i p \triangleright \mathtt{TBA}_i p : \varphi \triangleright p \in \mathit{deff}(a)\} \cup \\
&\{\varphi \wedge \mathtt{MBA}_i p \wedge \mathtt{MBA}_j \mathtt{TBA}_i p \triangleright \mathtt{TBA}_j \mathtt{TBA}_i p : \\
&\qquad\qquad \varphi \triangleright p \in \mathit{deff}(a), i \neq j\} \cup \\
&\{\varphi \wedge \neg\mathtt{MBA}_i p \wedge \mathtt{FBA}_j \mathtt{MBA}_i p \wedge \mathtt{OBS}_j p \triangleright \mathtt{TBA}_j \mathtt{TBA}_i p : \\
&\qquad\qquad \varphi \triangleright p \in \mathit{deff}(a), i \neq j\}.
\end{aligned}$$

6 Epistemic Actions

Many kinds of epistemic actions exist, in particular sensing actions and actions of communication between agents such as informative and interrogative speech acts. We here only consider the two kinds that we need in order to account for false-belief tasks.

1. A group of agents $J \subseteq \mathbb{A}$ starts to observe a propositional variable p. We focus on two versions, according to whether the agents in J observe the other agents starting to observe or not. When they don't we have the first-order $\mathtt{startobs}^1(J, p)$; and when everybody observes the others starting their observation we have the second-order $\mathtt{startobs}^2(J, p)$. The former actually reduces to a sequence of $\mathtt{startobs}^1(i, p)$, for any order of the $i \in J$.
2. An agent $i \in \mathbb{A}$ stops observing either p or another agent j's observation of p. The former is denoted by $\mathtt{stopobs}(i, p)$ and the latter by $\mathtt{stopobs}(i, j, p)$.

We stress that if one of the members of J has a false belief about p then $\mathtt{startobs}^1(J, p)$ requires the revision of i's beliefs about p. Similarly, $\mathtt{startobs}^2(J, p)$ requires the revision of false beliefs about another members' belief about p. We now define direct effects and derive indirect effects.

Starting Individual Observation

The action of agent $i \in \mathbb{A}$ starting to observe a propositional variable p without learning about others' belief change is denoted by $a = \mathtt{startobs}^1(i, p)$. Its direct effect is that $\mathtt{OBS}_i p = \mathtt{TBA}_i p \wedge \neg\mathtt{MBA}_i p$ becomes true:

$$\mathit{deff}(a) = \{\neg\mathtt{TBA}_i p \triangleright \mathtt{TBA}_i p, \mathtt{MBA}_i p \triangleright \mathtt{MBA}_i p\}.$$

As for the indirect effects, applying Principle (M) to the direct effects we derive the following:

$$\begin{aligned}
&\mathit{ieff}(a) = \\
&\{\neg\mathtt{TBA}_i p \wedge \mathtt{MBA}_j \mathtt{TBA}_i p \ \triangleright\ \mathtt{TBA}_j \mathtt{TBA}_i p : j \in \mathbb{A}, j \neq i\} \cup \\
&\{\mathtt{MBA}_i p \wedge \mathtt{MBA}_j \mathtt{MBA}_i p \ \triangleright\ \mathtt{TBA}_j \mathtt{MBA}_i p : j \in \mathbb{A}, j \neq i\}.
\end{aligned}$$

Even if we have already motivated Principle (M) in Section 4, let us explain here why these conditional effects are intuitively correct. For the first subset, suppose $\mathtt{MBA}_j \mathtt{TBA}_i p$ is true; then j does not observe whether i has a true belief about p but holds a belief about it (that is, about $\mathtt{TBA}_i p$); and j will keep that belief because i's observation change is only observed by i. Suppose moreover that $\mathtt{TBA}_i p$ is false. As $\mathtt{startobs}^1(i, p)$ makes $\mathtt{TBA}_i p$ true, the truth status of j's belief about $\mathtt{TBA}_i p$ will be flipped: either from $\mathtt{TBA}_j \mathtt{TBA}_i p$ to $\neg\mathtt{TBA}_j \mathtt{TBA}_i p$, or from $\neg\mathtt{TBA}_j \mathtt{TBA}_i p$ to $\mathtt{TBA}_j \mathtt{TBA}_i p$. The second subset can be motivated in a similar way.

These two indirect effects of $\mathtt{startobs}^1(i, p)$ on $\mathtt{TBA}_j \mathtt{TBA}_i p$ and $\mathtt{TBA}_j \mathtt{MBA}_i p$ are the only ones: the truth values of all remaining atoms $\mathtt{MBA}_j \mathtt{TBA}_i p$ and $\mathtt{MBA}_j \mathtt{MBA}_i p$ are stable because $\mathtt{startobs}^1(i, p)$ only modifies the status of the first-order $\mathtt{MBA}_i p$, but not of second-order mere beliefs.

Starting Shared Observation

The action of agents $J \subseteq \mathbb{A}$ starting to observe a propositional variable p while learning that the other agents in J

$$\begin{tikzcd}
\boxed{\begin{matrix} p, \\ \mathtt{OBS}_\mathtt{S}p, \\ \mathtt{OBS}_\mathtt{A}p, \\ \mathtt{OBS}_\mathtt{A}\mathtt{TBA}_\mathtt{S}p, \\ \mathtt{OBS}_\mathtt{A}\mathtt{MBA}_\mathtt{S}p \end{matrix}} \arrow[r, "{\mathtt{stopobs}(\mathtt{S},p)}"] & \boxed{\begin{matrix} p, \\ \mathtt{LBA}_\mathtt{S}p, \\ \mathtt{OBS}_\mathtt{A}p, \\ \mathtt{OBS}_\mathtt{A}\mathtt{TBA}_\mathtt{S}p, \\ \mathtt{OBS}_\mathtt{A}\mathtt{MBA}_\mathtt{S}p \end{matrix}} \arrow[r, "{\mathtt{stopobs}(\mathtt{A},\mathtt{S},p)}"] & \boxed{\begin{matrix} p, \\ \mathtt{LBA}_\mathtt{S}p, \\ \mathtt{OBS}_\mathtt{A}p, \\ \mathtt{LBA}_\mathtt{A}\mathtt{TBA}_\mathtt{S}p, \\ \mathtt{LBA}_\mathtt{A}\mathtt{MBA}_\mathtt{S}p \end{matrix}} \arrow[r, "{\mathtt{startobs}^1(\mathtt{S},p)}"] & \boxed{\begin{matrix} p, \\ \mathtt{OBS}_\mathtt{S}p, \\ \mathtt{OBS}_\mathtt{A}p, \\ \mathtt{LBA}_\mathtt{A}\mathtt{TBA}_\mathtt{S}p, \\ \mathtt{FBA}_\mathtt{A}\mathtt{MBA}_\mathtt{S}p \end{matrix}} \arrow[r, "{\mathtt{flip}(p)}"] & \boxed{\begin{matrix} \neg p, \\ \mathtt{OBS}_\mathtt{S}p, \\ \mathtt{OBS}_\mathtt{A}p, \\ \mathtt{FBA}_\mathtt{A}\mathtt{TBA}_\mathtt{S}p, \\ \mathtt{FBA}_\mathtt{A}\mathtt{MBA}_\mathtt{S}p \end{matrix}}
\end{tikzcd}$$

Figure 2: The first variant of the Sally-Anne Task of Section 7.

also start observing p is noted $a = \mathtt{startobs}^2(J,p)$. Its direct effects are:

$$\begin{aligned} \mathit{deff}(a) = {} & \{\neg\mathtt{TBA}_i p \triangleright \mathtt{TBA}_i p : i \in J\} \cup \\ & \{\neg\mathtt{TBA}_j\mathtt{TBA}_i p \triangleright \mathtt{TBA}_j\mathtt{TBA}_i p : i,j \in J, j \neq i\} \cup \\ & \{\neg\mathtt{TBA}_j\mathtt{MBA}_i p \triangleright \mathtt{TBA}_j\mathtt{MBA}_i p : i,j \in J, j \neq i\} \cup \\ & \{\mathtt{MBA}_i p \triangleright \mathtt{MBA}_i p : i \in J\} \cup \\ & \{\mathtt{MBA}_j\mathtt{TBA}_i p \triangleright \mathtt{MBA}_j\mathtt{TBA}_i p : i,j \in J, j \neq i\} \cup \\ & \{\mathtt{MBA}_j\mathtt{MBA}_i p \triangleright \mathtt{MBA}_j\mathtt{MBA}_i p : i,j \in J, j \neq i\}. \end{aligned}$$

The direct effects guarantee that $\mathtt{startobs}^2(J,p)$ makes $\mathtt{OBS}_i p \wedge \mathtt{OBS}_i\mathtt{TBA}_j p \wedge \mathtt{OBS}_i\mathtt{MBA}_j p$ true for all i, j in J.

As for the indirect effects of $\mathtt{startobs}^2(J,p)$, Principle (M) applies to $\neg\mathtt{TBA}_i p \triangleright \mathtt{TBA}_i p$ and $\mathtt{MBA}_i p \triangleright \mathtt{MBA}_i p$ and results in the following effects:

$$\begin{aligned} & \mathit{ieff}(a) = \\ & \{\neg\mathtt{TBA}_i p \wedge \mathtt{MBA}_j\mathtt{TBA}_i p \triangleright \mathtt{TBA}_j\mathtt{TBA}_i p : i \in J, j \in \mathbb{A} \setminus J\} \cup \\ & \{\mathtt{MBA}_i p \wedge \mathtt{MBA}_j\mathtt{MBA}_i p \triangleright \mathtt{TBA}_j\mathtt{MBA}_i p : i \in J, j \in \mathbb{A} \setminus J\}. \end{aligned}$$

Note that this only applies to $j \notin J$: for $i, j \in J$, both $\mathtt{MBA}_j\mathtt{TBA}_i p$ and $\mathtt{MBA}_j\mathtt{MBA}_i p$ are made false by the direct effects. There are no other effects as the beliefs of agents not in J as well as their higher-order mere beliefs are not changed.

Ceasing to Observe a Fact of the World

The action $a = \mathtt{stopobs}(i,p)$ of agent $i \in \mathbb{A}$ ceasing to observe $p \in \mathbb{P}$ has a single direct effect:

$$\mathit{deff}(a) = \{\mathtt{TBA}_i p \wedge \neg\mathtt{MBA}_i p \triangleright \mathtt{MBA}_i p\}.$$

This guarantees that $\mathtt{stopobs}(i,p)$ makes $\mathtt{MBA}_i p$ true; that is, i no longer observes α. The role of the condition $\mathtt{TBA}_i p$ is to exclude the case where i has no belief about p: then $\mathtt{MBA}_i p$ should not be flipped, that is, i remains ignorant about p.

Just like the direct effects, the indirect effects apply only when i observes p: if moreover j has a belief about $\mathtt{MBA}_i p$ but does not observe it then j does not observe any change of $\mathtt{MBA}_i p$; and as $a = \mathtt{stopobs}(i,p)$ changes $\mathtt{MBA}_i p$, the truth value of $\mathtt{TBA}_j\mathtt{MBA}_i p$ gets flipped. We therefore have:

$$\begin{aligned} & \mathit{ieff}(a) = \\ & \{\mathtt{TBA}_i p \wedge \neg\mathtt{MBA}_i p \wedge \mathtt{MBA}_j\mathtt{MBA}_i p \triangleright \mathtt{TBA}_j\mathtt{MBA}_i p : j \neq i\}. \end{aligned}$$

These are the only indirect effects of $a = \mathtt{stopobs}(i,p)$: the other second-order atoms remain stable. In particular, $\mathtt{TBA}_j\mathtt{TBA}_i p$, $\mathtt{TBA}_j\mathtt{MBA}_i p$, $\mathtt{MBA}_j\mathtt{TBA}_i p$, $\mathtt{MBA}_j\mathtt{MBA}_i p$ remain unchanged, reflecting that $\mathtt{stopobs}(i,p)$ is public for the agents observing $\mathtt{TBA}_i p$ and $\mathtt{MBA}_i p$.

Ceasing to Observe Another Agent

The action $a = \mathtt{stopobs}(i,j,p)$ of agent $i \in \mathbb{A}$ ceasing to observe whether j observes p has two direct effects:

$$\begin{aligned} \mathit{deff}(a) = \{ & \mathtt{OBS}_i\mathtt{TBA}_j p \wedge \mathtt{OBS}_i\mathtt{MBA}_j p \triangleright \mathtt{MBA}_i\mathtt{TBA}_j p, \\ & \mathtt{OBS}_i\mathtt{TBA}_j p \wedge \mathtt{OBS}_i\mathtt{MBA}_j p \triangleright \mathtt{MBA}_i\mathtt{MBA}_j p\}. \end{aligned}$$

This guarantees that $\mathtt{stopobs}(i,p)$ makes $\mathtt{MBA}_i\mathtt{TBA}_j p \wedge \mathtt{MBA}_i\mathtt{MBA}_j p$ true; that is, i no longer observes whether j observes p. Indirect effects of $\mathtt{stopobs}(i,j,p)$ would be of modal depth three and are therefore not considered here.

7 Second-Order False-Belief Tasks

We now demonstrate the usefulness of our action repertoire by applying it to false-belief tasks. The first example involves second-order beliefs and is analogous to the second-order chocolate task (Flobbe et al. 2008; Arslan, Taatgen, and Verbrugge 2013) in Bolander's version (2018). The only difference with the simple Sally-Anne Task is that after leaving the room Sally peeks through the window. This allows her to observe the position of the marble again without Anne noticing. The question is where Anne believes that Sally believes the marble is. Figure 2 shows the evolution of *REDA* states. For brevity, we model Anne's beliefs about Sally's beliefs, but not vice versa. We also make use of our abbreviations $\mathtt{OBS}$, $\mathtt{LBA}$, $\mathtt{FBA}$, and $\mathtt{NBA}$ in order to compactly represent the various combinations of $\mathtt{TBA}$ and $\mathtt{MBA}$.

We model Sally leaving the room as two successive actions: first, Sally stops observing the marble (she turns away and walks through the door), and then Anne stops observing Sally (once Sally is gone, Anne no longer sees her). The first action $\mathtt{stopobs}(\mathtt{S},p)$ makes $\mathtt{MBA}_\mathtt{S}p$ true: Sally's observation of p is turned into a lucky belief. As the condition $\mathtt{MBA}_\mathtt{A}\mathtt{MBA}_\mathtt{S}p$ is false, Anne's second-order beliefs do not change. The second action $\mathtt{stopobs}(\mathtt{A},\mathtt{S},p)$ further turns Anne's observations of $\mathtt{TBA}_\mathtt{S}p$ and $\mathtt{MBA}_\mathtt{S}p$ into lucky beliefs.

We then model Sally peeking through the window as the action $\mathtt{startobs}^1(\mathtt{S},p)$. As a direct effect, only $\mathtt{MBA}_\mathtt{S}p$ is flipped, making $\mathtt{OBS}_\mathtt{S}p$ true. As for indirect effects, only $\mathtt{TBA}_\mathtt{A}\mathtt{MBA}_\mathtt{S}p$ is flipped, turning the lucky belief into a false belief. In the resulting state, Anne now has the false belief that Sally's belief is a mere belief.

Finally, the action of removing the marble from the basket is simply the ontic action $\mathtt{flip}(p)$. Besides the direct effect, only the following indirect effect applies:

$$\neg\mathtt{MBA}_\mathtt{S}p \wedge \mathtt{FBA}_\mathtt{A}\mathtt{MBA}_\mathtt{S}p \wedge \mathtt{OBS}_\mathtt{A}p \triangleright \mathtt{TBA}_\mathtt{A}\mathtt{TBA}_\mathtt{S}p.$$

In the end, the marble is in the box, and both agents can observe this. However Anne has false beliefs about both Sally's true and mere beliefs about p: she falsely believes that $\neg\mathtt{TBA}_\mathtt{S}p$ and $\mathtt{MBA}_\mathtt{S}p$, i.e., she falsely believes that $\mathtt{FBA}_\mathtt{S}p$.

In the second version we switch the final $\texttt{flip}(p)$ and $\texttt{startobs}^1(\texttt{S}, p)$. In this case, Sally first obtains a false belief about p as result of the $\texttt{flip}(p)$ action. One effect of $\texttt{startobs}^1(\texttt{S}, p)$ is then that Sally revises her belief. The resulting state is the same as the final state in Figure 2.

In the third version, Anne eventually notices Sally looking in. We add a final $\texttt{startobs}^2(\{\texttt{A}, \texttt{S}\}, p)$ (assuming we also model Sally's second order beliefs), resulting in Anne revising her second-order false belief about Sally's first order false belief.

8 Epistemic-Doxastic Planning

We now sketch a planning formalism using our logic.

Similar to classical planning, a *REDA planning task* consists of an *initial state* $I \subseteq$ *REDA*$^{\leq 2}$, a set of *actions* A, and a *goal formula* $\gamma \in \mathcal{L}_{\mathsf{bool}}(\textit{REDA})$. Each action $a \in A$ is an action from our repertoire (i.e., an arbitrary ontic action or one of the epistemic $\texttt{startobs}$, $\texttt{stopobs}$ actions), annotated with an additional *precondition* $\textit{pre}(a) \in \mathcal{L}_{\mathsf{bool}}(\textit{REDA})$ that specifies the condition under which a is applicable.

A *plan* is then a sequence of actions $a_1, \ldots, a_n \in A$ such that $a_1, \ldots, a_n$ are sequentially applicable from I and the resulting final state satisfies γ. The plan existence problem is the problem of deciding, for a given planning task, whether there exists a plan. False-belief tasks can be turned into such planning tasks; for example, one may wish to know whether there exists a plan that produces a false belief about p of some agent i while keeping the beliefs of j correct.

REDA planning can be polynomially reduced to classical planning with conditional effects and vice versa: When translating from *REDA* to classical planning, we directly use the atoms occurring in the *REDA* task as state variables. When translating from classical to *REDA* planning, we use $\mathbb{P}$ for the state variables and never need any atom from *REDA*$^{\geq 1}$. The translations between our flip-based and classical add/delete-based conditional effects are straightforward. Since classical plan existence with conditional effects and propositional preconditions is PSPACE-complete (Bylander 1994; Nebel 2000), we obtain the following result:

Theorem 2 *REDA plan existence is* PSPACE*-complete.*

9 Discussion and Conclusion

We have introduced a new action formalism that can be used to model change of agents' knowledge and beliefs and that encompasses belief revision. It is based on the concepts of observability of propositional facts and of other agents' first-order beliefs. We have illustrated how our formalism can model false-belief tasks such as the Sally-Anne Task and its second-order version. We proved that reasoning reduces to classical propositional logic. This is due to the restriction that epistemic atoms are repetition-free. The same restriction was exploited under the denomination 'agent-alternating formulas' by Ding, Holliday, and Zhang (2019) in order to reduce reasoning with introspection to reasoning without introspection (though on 'knowledge-that' and 'belief-that' operators). Finally, we have shown how our formalism can be used for epistemic planning with the same complexity as classical planning.

There exist other epistemic action languages that can be used for similar purposes. Bolander (2018) describes a simple action language with the explicit purpose of modelling higher-order false-belief tasks, which translates directly into dynamic epistemic logic with edge-conditioned update models. Baral et al. (2022) introduce $m\mathcal{A}^*$, a general-purpose action language that includes ontic actions, announcements, and sensing actions where agents can have different levels of observation. The Sally-Anne Task can be formalised in the original version of $m\mathcal{A}^*$, but not higher-order false-belief tasks. This is due to the inability of the semantics to deal with false beliefs about observations. In a follow-up paper, Pham et al. (2022) introduced a new semantics for $m\mathcal{A}^*$, also based on edge-conditioned update models, that correctly handles such situations. Another approach has been proposed by Lorini and Romero (2019), which is based on belief bases. Similar to the original version of $m\mathcal{A}^*$, it does not support false beliefs about observations and thus can only model first-order false-belief tasks. Wan, Fang, and Liu (2021) propose another epistemic planning formalism based on KD45 update and revision.

Those approaches differ from ours in the way they model observability. The papers by Baral et al., Pham et al., and Wan et al. consider the observability of actions (agent i observes an action taking place), while Bolander and Lorini and Romero consider the observability of agents (agent i observes agent j). In contrast to our approach, which considers the observability of epistemic-doxastic atoms, these approaches model only beliefs and not knowledge.

On the other hand, a problem with most DEL-based approaches is that they do not account for belief revision. For example, both $m\mathcal{A}^*$ and Bolander's approach assume KD45 models in their semantics. However, KD45 models are not closed under product update: if an agent has a false belief about φ and senses whether φ is the case then the updated model is no longer a KD45 model (Balbiani et al. 2012; Herzig 2017). In the literature this limitation has been overcome by moving from KD45 models to plausibility orderings (van Ditmarsch 2005; Baltag and Smets 2006; Aucher 2008; Andersen, Bolander, and Jensen 2015). Our approach does not require such complex devices, as illustrated by the second example from Section 7: in a state where $\texttt{FBA}_i p$ is true, the execution of the action $\texttt{startobs}^1(i, p)$ results in a state where $\texttt{OBS}_i p$ and thus $\texttt{TBA}_i p$ hold.

As a drawback, there are situations that we cannot model in our formalism. For example, consider a state where agent j observes *that* agent i *does not* observe p. For this to be true, it must be the case that $\texttt{OBS}_j\texttt{TBA}_i p$ and $\texttt{OBS}_j\texttt{MBA}_i p$, as in any other case j would be uncertain about the observation status of i. However, this assumption seems too strong in general: since j can now observe the value of $\texttt{TBA}_i p$ and $\texttt{MBA}_i p$ separately, she also knows whether i has a lucky belief, a false belief, or no belief about p. This means that using combinations of *REDA* atoms, it is impossible to model the situation where i observes that j does not observe p, and at the same time i has no knowledge about j's beliefs. We have also so far excluded beliefs of depth greater than two. In future work, we plan to lift these limitations by considering richer fragments of epistemic-doxastic logic.

Acknowledgments

This work was supported by the ANR LabEx CIMI (grant ANR-11-LABX-0040) within the French State Programme "Investissements d'Avenir", the AI Chair BE4musIA of the French National Research Agency (ANR-20-CHIA-0028), and the EU ICT-48 2020 project TAILOR (No. 952215).

References

Andersen, M. B.; Bolander, T.; and Jensen, M. H. 2015. Don't plan for the unexpected: Planning based on plausibility models. *Logique et Analyse*, 145–176.

Arslan, B.; Taatgen, N.; and Verbrugge, R. 2013. Modeling developmental transitions in reasoning about false beliefs of others. In *Proc. ICCM 2013*, 77–82.

Aucher, G. 2008. Internal models and private multi-agent belief revision. In *Proc. AAMAS 2008*, 721–727.

Aucher, G. 2015. Intricate Axioms as Interaction Axioms. *Studia Logica*, 103(5): 1035–1062.

Aucher, G.; and Bolander, T. 2013. Undecidability in epistemic planning. In *Proc. IJCAI 2013*, 27–33.

Balbiani, P.; van Ditmarsch, H.; Herzig, A.; and de Lima, T. 2012. Some Truths Are Best Left Unsaid. In *Advances in Modal Logic 9*, 36–54.

Baltag, A.; and Smets, S. 2006. Dynamic Belief Revision over Multi-Agent Plausibility Models. In *Proc. LOFT 2006*.

Baral, C.; Gelfond, G.; Pontelli, E.; and Son, T. C. 2022. An action language for multi-agent domains. *Artif. Intell.*, 302: 103601.

Baron-Cohen, S.; Leslie, A. M.; and Frith, U. 1985. Does the autistic child have a theory of mind? *Cognition*, 21(1): 37–46.

Bolander, T. 2018. Seeing Is Believing: Formalising False-Belief Tasks in Dynamic Epistemic Logic. In *Jaakko Hintikka on Knowledge and Game-Theoretical Semantics*, 207–236. Springer International Publishing.

Bolander, T.; and Andersen, M. B. 2011. Epistemic planning for single and multi-agent systems. *Journal of Applied Non-Classical Logics*, 21(1): 9–34.

Bolander, T.; Charrier, T.; Pinchinat, S.; and Schwarzentruber, F. 2020. DEL-based epistemic planning: Decidability and complexity. *Artif. Intell.*, 287: 103304.

Bolander, T.; Engesser, T.; Mattmüller, R.; and Nebel, B. 2018. Better Eager Than Lazy? How Agent Types Impact the Successfulness of Implicit Coordination. In *Proc. KR 2018*, 445–453.

Bylander, T. 1994. The computational complexity of propositional STRIPS planning. *Artificial Intelligence*, 69: 165–204.

Chignell, A. 2018. The Ethics of Belief. In *The Stanford Encyclopedia of Philosophy*. Stanford University, Spring 2018 edition.

Cooper, M. C.; Herzig, A.; Maffre, F.; Maris, F.; Perrotin, E.; and Régnier, P. 2021. A lightweight epistemic logic and its application to planning. *Artif. Intell.*, 298: 103437.

Ding, Y.; Holliday, W. H.; and Zhang, C. 2019. When Do Introspection Axioms Matter for Multi-Agent Epistemic Reasoning? In *Proc. TARK 2019*, 121–139.

Dissing, L.; and Bolander, T. 2020. Implementing Theory of Mind on a Robot Using Dynamic Epistemic Logic. In *Proc. IJCAI 2020*, 1615–1621.

Engesser, T.; Bolander, T.; Mattmüller, R.; and Nebel, B. 2017. Cooperative Epistemic Multi-Agent Planning for Implicit Coordination. In *Proc. M4M@ICLA 2017*, 75–90.

Engesser, T.; and Miller, T. 2020. Implicit Coordination Using FOND Planning. In *Proc. AAAI 2020*, 7151–7159.

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

Flobbe, L.; Verbrugge, R.; Hendriks, P.; and Krämer, I. 2008. Children's Application of Theory of Mind in Reasoning and Language. *J. Log. Lang. Inf.*, 17(4): 417–442.

Herzig, A. 2017. Dynamic epistemic logics: promises, problems, shortcomings, and perspectives. *J. Appl. Non Class. Logics*, 27(3-4): 328–341.

Herzig, A.; Lorini, E.; and Maffre, F. 2018. Possible Worlds Semantics Based on Observation and Communication. In *Jaakko Hintikka on Knowledge and Game-Theoretical Semantics*, 339–362. Springer International Publishing.

Herzig, A.; and Perrotin, E. 2021. True Belief and Mere Belief About a Proposition and the Classification of Epistemic-Doxastic Situations. *Filosofiska Notiser*, 8(1): 103–117.

Ichikawa, J. J.; and Steup, M. 2018. The Analysis of Knowledge. In *The Stanford Encyclopedia of Philosophy*. Stanford University, Summer 2018 edition.

Kominis, F.; and Geffner, H. 2015. Beliefs in multiagent planning: from one agent to many. In *Proc. ICAPS 2015*, 147–155.

Lakemeyer, G.; and Lespérance, Y. 2012. Efficient Reasoning in Multiagent Epistemic Logics. In *Proc. ECAI 2012*, 498–503.

Lorini, E.; and Romero, F. 2019. Decision Procedures for Epistemic Logic Exploiting Belief Bases. In *Proc. AAMAS 2019*, 944–952.

Muise, C.; Belle, V.; Felli, P.; McIlraith, S. A.; Miller, T.; Pearce, A. R.; and Sonenberg, L. 2022. Efficient multi-agent epistemic planning: Teaching planners about nested belief. *Artif. Intell.*, 302: 103605.

Nebel, B. 2000. On the Compilability and Expressive Power of Propositional Planning Formalisms. *J. Artif. Intell. Res.*, 12: 271–315.

Pham, L.; Izmirlioglu, Y.; Son, T. C.; and Pontelli, E. 2022. A New Semantics for Action Language $m\mathcal{A}^*$. In *Proc. PRIMA 2022*, 553–562.

van Ditmarsch, H. P. 2005. Prolegomena to Dynamic Logic for Belief Revision. *Synth.*, 147(2): 229–275.

Voorbraak, F. 1993. *As Far as I Know. Epistemic Logic and Uncertainty*. Ph.D. thesis, Universiteit Utrecht.

Wan, H.; Fang, B.; and Liu, Y. 2021. A general multi-agent epistemic planner based on higher-order belief change. *Artif. Intell.*, 301: 103562.

Wimmer, H.; and Perner, J. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1): 103–128.