Planning Over Multi-Agent Epistemic States: A Classical Planning Approach

Citation for published version:
Muise, C, Belle, V, Felli, P, McIlraith, S, Miller, T, Pearce, AR & Sonenberg, L 2015, Planning Over Multi-Agent Epistemic States: A Classical Planning Approach. in *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence.* AAAI Press, pp. 3327-3334, Twenty-Ninth AAAI Conference on Artificial Intelligence, 25/01/15. <http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9974>

Published In:
Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence

Planning Over Multi-Agent Epistemic States: A Classical Planning Approach

Christian Muise[*], Vaishak Belle[†], Paolo Felli[*], Sheila McIlraith[†]
Tim Miller[*], Adrian R. Pearce[*], Liz Sonenberg[*]

[*]Department of Computing and Information Systems, University of Melbourne
[†]Department of Computer Science, University of Toronto
[*]{christian.muise,paolo.felli,tmiller,adrianrp,l.sonenberg}@unimelb.edu.au, [†]{vaishak,sheila}@cs.toronto.edu

Abstract

Many AI applications involve the interaction of multiple autonomous agents, requiring those agents to reason about their own beliefs, as well as those of other agents. However, planning involving nested beliefs is known to be computationally challenging. In this work, we address the task of synthesizing plans that necessitate reasoning about the beliefs of other agents. We plan from the perspective of a single agent with the potential for goals and actions that involve nested beliefs, non-homogeneous agents, co-present observations, and the ability for one agent to reason *as if* it were another. We formally characterize our notion of planning with nested belief, and subsequently demonstrate how to automatically convert such problems into problems that appeal to classical planning technology. Our approach represents an important first step towards applying the well-established field of automated planning to the challenging task of planning involving nested beliefs of multiple agents.

1 Introduction

AI applications increasingly involve the interaction of multiple agents – be they intelligent user interfaces that interact with human users, gaming systems, or multiple autonomous robots interacting together in a factory setting. In the absence of prescribed coordination, it is often necessary for individual agents to synthesize their own plans, taking into account not only their own capabilities and beliefs about the world but also their beliefs about other agents, including what each of the agents will come to believe as the consequence of the actions of others. To illustrate, consider the scenario where Larry and Moe plan to work together on an assembly task. Each knows what needs to be done and can plan accordingly. Unbeknownst to Moe, Larry decides to start the job early. Larry believes that Moe believes that assembly has not yet commenced. As a consequence, Larry's plan must include a communication action to inform Moe of the status of the assembly when Moe arrives.

In this paper, we examine the problem of synthesizing plans in such settings. In particular, given a finite set of agents, each with: (1) (possibly incomplete and incorrect) beliefs about the world and about the beliefs of other agents; and (2) differing capabilities including the ability to perform actions whose outcomes are unknown to other agents; we are interested in synthesizing a plan to achieve a goal condition. Planning is at the belief level and as such, while we consider the execution of actions that can change the state of the world (ontic actions) as well as an agent's state of knowledge or belief (epistemic or more accurately doxastic actions, including communication actions), all outcomes are with respect to belief. Further, those beliefs respect the KD45$_n$ axioms of epistemic logic (Fagin et al. 1995). Finally, we take a *perspectival view*, planning from the viewpoint of a single agent. We contrast this with traditional multi-agent planning which generates a coordinated plan to be executed by multiple agents (e.g., (Brenner and Nebel 2009)).

We focus on computational aspects of this synthesis task, leaving exploration of interesting theoretical properties to a companion paper. To this end, we propose a means of encoding a compelling but restricted subclass of our synthesis task as a classical planning problem, enabling us to exploit state-of-the-art classical planning techniques to synthesize plans for these challenging planning problems. Our approach relies on two key restrictions: (1) we do not allow for disjunctive belief; and (2) the depth of nested belief is bounded. A key aspect of our encoding is the use of *ancillary conditional effects* – additional conditional effects of actions which enforce desirable properties such as epistemic modal logic axioms (cf. Section 3), and allow domain modellers to encode conditions under which agents are mutually aware of actions (cf. Section 4). By encoding modal logic axioms as effects of actions, we are using our planner to perform epistemic reasoning. As such, our planning machinery additionally supports answering queries involving the nested beliefs of agents (cf. Section 5): e.g, "Does Agent 1 believe that Agent 2 believes they can achieve the goal?".

Computational machinery for epistemic reasoning has historically appealed to theorem proving or model checking (e.g., (van Eijck 2004)), while epistemic planning, recently popularized within the *Dynamic Epistemic Logic* (DEL) community, has largely focused on theoretical concerns (e.g., (Löwe, Pacuit, and Witzel 2011)). The work presented here is an important first step towards leveraging state-of-the-art planning technology to address rich epistemic planning problems of the sort examined by the DEL community. Indeed, we can readily solve existing examples in the DEL literature (cf. Section 6). We further discuss the relationship of our work to other work in epistemic reasoning and planning at the end of this paper.

Example 1 (Grapevine). We will use a common example to explain the concepts introduced throughout the paper. Consider a scenario where a group of agents each have their own secret to (possibly) share with one another. Each agent can move freely between a pair of rooms, and broadcast any secret they currently believe to everyone in the room. Initially they only believe their own unique secret. Goals we might pose include the universal spread of information (everyone believes every secret), misconception (an agent holds a false belief about someone else's belief), etc. We will use $1, 2, \cdots$ to represent the agents, and $s_1, s_2, \cdots$ to represent their secrets, respectively.

2 Specification

The general aim of this work is to address problems similar to DEL planning (Bolander and Andersen 2011) using the computational machinery of automated planning. We use DEL to formally specify our planning system. Our presentation below is terse, and we refer interested readers to van Ditmarsch, van der Hoek, and Kooi (2007) for a more comprehensive overview.

Let $\mathcal{P}$, $\mathcal{A}$, and Ag respectively be finite sets of propositions, actions, and agents. The set of well-formed formulae, $\mathcal{L}$, for DEL is obtained from the following grammar:

$$\phi ::= p \mid \phi \wedge \phi' \mid B_i\phi \mid [\alpha]\phi \mid \neg\phi$$

in which $p \in \mathcal{P}$, $\alpha \in \mathcal{A}$, and $i \in Ag$. $B_i\phi$ should be interpreted as "agent i believes ϕ." The semantics is given using *Kripke structures* (Fagin et al. 1995). Given a world w, standing for some state of affairs, such a structure determines (by means of an *accessibility relation*) the worlds that an agent considers possible when at w. (That is, the agent is unsure which world it is truly in). A *model* M is the set of all worlds, an accessibility relation between these worlds for each agent i, and a function specifying which propositions are true in each world. Informally, the meaning of formulas wrt a pair (M, w) is as follows: p holds if it is true in w, $\phi \wedge \psi$ holds if both ϕ and ψ hold, $\neg\phi$ holds if ϕ does not hold at (M, w), $B_i\phi$ if ϕ holds in all worlds agent i considers possible at w, and $[\alpha]\phi$ holds if ϕ holds after *applying* action α to (M, w). The semantics is defined formally in terms of $\models$, where $M, w \models \phi$ means that ϕ holds in world w for model M.

As discussed by Fagin et al. (1995), constraints on Kripke structures lead to particular properties of belief. If the Kripke structure is *serial*, *transitive*, and *Euclidean* we obtain (arguably) the most common properties of belief:

K	$B_i\phi \wedge B_i(\phi \supset \psi) \supset B_i\psi$	(Distribution)
D	$B_i\phi \supset \neg B_i\neg\phi$	(Consistency)
4	$B_i\phi \supset B_iB_i\phi$	(Positive introspection)
5	$\neg B_i\phi \supset B_i\neg B_i\phi$	(Negative introspection)

These axioms collectively form the system referred to as *KD45$_n$*, where n specifies that there are multiple agents in the environment. From the axioms, additional theorems can be derived. For example, in this work, we use the following theorems for reducing neighbouring belief modalities involving the same agent into a single belief modality:

$$B_iB_i\phi \equiv B_i\phi \qquad B_i\neg B_i\phi \equiv \neg B_i\phi$$
$$\neg B_i\neg B_i\phi \equiv B_i\phi \qquad \neg B_iB_i\phi \equiv \neg B_i\phi$$

We can now define a planning problem as follows:

Definition 1. Multi-Agent Epistemic Planning Problem

A ***m**ulti-agent **e**pistemic **p**lanning* (MEP) problem $\mathcal{D}$ is a tuple of the form $\langle \mathcal{P}, \mathcal{A}, Ag, \mathcal{I}, \mathcal{G} \rangle$, where $\mathcal{P}$, $\mathcal{A}$, and Ag are as above, $\mathcal{I}$ is the *initial theory*, and $\mathcal{G}$ is the *goal condition*. Each $a \in \mathcal{A}$ is assumed to be of the form $\langle \pi, \{(\gamma_1, l_1), \ldots, (\gamma_k, l_k)\} \rangle$, where π is called the *precondition* of a, γ_i is called the *condition* of a *conditional effect*, and l_i is called the *effect* of a conditional effect. Finally, we assume $\mathcal{G}$, $\mathcal{I}$, π, γ_i, and l_i are all well-formed formulae over $\mathcal{L}$, excluding the $[\alpha]$ modality.

Following Reiter (2001) and van Ditmarsch, van der Hoek, and Kooi (2007), the above action formalization can be expressed as standard *precondition* and *successor state axioms*, which would then define the meaning of $[\alpha]\phi$ in DEL. By extension, we say that given a domain $\mathcal{D} = \langle \mathcal{P}, \mathcal{A}, Ag, \mathcal{I}, \mathcal{G} \rangle$, the sequence of actions $a_1, \ldots, a_k$ *achieves* $\mathcal{G}$ iff for any (M, w) such that $M, w \models \mathcal{I}$, we have $M, w \models [a_1] \ldots [a_k]\mathcal{G}$. Thus, the *plan synthesis task* is one of finding a sequence of actions $\vec{a}$ that achieves the goal condition $\mathcal{G}$.

Not surprisingly, reasoning (and planning) in these logical frameworks is computationally challenging (Fagin et al. 1995; Aucher and Bolander 2013). In this work, we limit our attention to a planning framework described using a fragment of epistemic logic. First, we consider reasoning from the perspective of a single *root agent*; this is a perspectival view of the world. Second, we do not allow disjunctive formulae as a belief. Following Lakemeyer and Lespérance (2012), we define a *restricted modal literal* (RML) as one obtained from the following grammar:

$$\phi ::= p \mid B_i\phi \mid \neg\phi$$

where $p \in \mathcal{P}$ and $i \in Ag$. The depth of an RML is defined as: $depth(p) = 0$ for $p \in \mathcal{P}$, $depth(\neg\phi) = depth(\phi)$ and $depth(B_i\phi) = 1+ depth(\phi)$. We will view a conjunction of RMLs equivalently as a set, and denote the set of all RMLs with bounded depth d for a group of agents Ag as $\mathcal{L}_{RML}^{Ag,d}$.

We define a ***r**estricted **p**erspectival* multi-agent epistemic planning problem (RP-MEP problem) for depth bound d and the root agent $\star \in Ag$ as a MEP problem with the additional restrictions that: (1) every RML is from the perspective of the root agent – i.e., it is from the following set:

$$\{B_\star\phi \mid \phi \in \mathcal{L}_{RML}^{Ag,d}\} \cup \{\neg B_\star\phi \mid \phi \in \mathcal{L}_{RML}^{Ag,d}\},$$

and (2) there is no disjunctive belief: the initial theory, goal specification, and every precondition are sets of positive RMLs (i.e., no negated belief), every effect is a single RML, and every effect condition is a set of RMLs.

We focus on the class of RP-MEP problems with an aim to extend our work to the more general class in the future. We address the planning problem from the view of an acting agent, where the designated root agent $\star$ is the one for which we plan. Intuitively, this means that conditional effects are formulated in the context of the root agent; e.g., we would have a conditional effect of the form $(\{B_\star\gamma\}, B_\star l)$ for action a in a RP-MEP problem to capture the fact that the root agent will believe l if it believed γ before a occurred.

This admits a rich class of planning problems; e.g., it is reasonable to assume that the root agent's view of the world differs from what a particular agent i believes, and so another conditional effect of a might be $(\{B_\star\gamma\}, B_\star B_i\neg l)$ – even though the root agent believes doing a would make l true if γ holds, the root agent believes that i will believe $\neg l$ if γ holds. In particular, this is easily shown to generalize a standard assumption in the literature (Liu and Wen 2011) that all agents hold the same view of what changes after actions occur.

In the next section, we show how restricted perspectival multi-agent epistemic planning problems can be represented as a classical planning problem, where the key insight is to encode reasoning features (such as deduction in KD45$_n$) as *ramifications* realized using ordinary planning operators.

3 A Classical Encoding

In this section, we present our model for planning with nested belief in a classical planning setting. We assume that the state of the world represents the mental model of a particular agent, perhaps an omniscient agent, that perceives an environment that includes all other agents. As a result, all reasoning is from the perspective of this single agent. The fluents that are true in a state correspond to the RMLs that the agent believes, while the fluents that are false correspond to the RMLs that the agent does not believe. Action execution, then, is predicated on the agent *believing* that the preconditions are satisfied. Similarly, the mental model of the agent is updated according to the effects of an action. Note that we do not need to enforce a separation of ontic and epistemic effects – the same action can update belief about propositions as well as RMLs. This is due to the interpretation that the state of the world represents the mental model of a given agent: every effect is epistemic in this sense.

The remainder of the section will proceed as follows:

1. We present the framework for our encoding of an RP-MEP problem into classical planning.
2. We specify how the state of the world is updated to maintain the deductive closure of the agent's belief.
3. We describe how to address the situation when an agent is uncertain if a conditional effect fires (e.g. due to lack of knowledge), and how the agent removes the corresponding beliefs from its knowledge base.

Items 2 and 3, in particular, enable the introspection capabilities and the change in beliefs after actions in standard epistemic frameworks (Aucher and Bolander 2013).

3.1 Encoding RP-MEP

We begin by providing a quick background on the classical planning formalism we use. A classical planning problem consists of a tuple $\langle F, I, G, O\rangle$, where F is a set of fluent atoms, I is a complete setting of the fluents describing the initial state, G is a set of fluents describing the goal condition, and O is a set of operators. A *state* s is a subset of the fluents F with the interpretation that atoms not in s are false. Every operator $o \in O$ is a tuple $\langle Pre_o, \mathit{eff}^+_o, \mathit{eff}^-_o\rangle$, and we say that o is applicable in s iff $Pre_o \subseteq s$. The set eff^+_o (resp. eff^-_o) contains conditional effects describing the fluent atoms that should be added (resp. removed) from the state when applying the operator. Finally, every conditional effect in eff^+_o or eff^-_o is of the form $(C \to l)$ where C is the condition for the effect and l is a fluent that is the result of the effect. The condition C consists of a tuple $\langle C^+, C^-\rangle$ where C^+ is the set of fluents that must hold and C^- the set of fluents that must not hold. A conditional effect $(\langle C^+, C^-\rangle \to l)$ *fires* in state s iff $C^+ \subseteq s$ and $C^- \cap s = \emptyset$. Assuming o is applicable in s, and $\mathit{eff}^+_o(s)$ (resp. $\mathit{eff}^-_o(s)$) are the positive (resp. negative) conditional effects that fire in state s, the state of the world s' after applying o is defined as follows:

$$\begin{aligned} s' = s \setminus \{l \mid (C \to l) \in \mathit{eff}^-_o(s)\} \\ \cup \{l \mid (C \to l) \in \mathit{eff}^+_o(s)\} \end{aligned}$$

Our account of classical planning mirrors the standard representation (see, for example, (Ghallab, Nau, and Traverso 2004)), with the exception that we make explicit the fluent atoms that are added, deleted, required to be in, or required to be absent from the state of the world. This simplifies the exposition when we encode nested beliefs as a classical planning problem. Intuitively, every RML in $\mathcal{L}^{Ag,d}_{RML}$ will correspond to a single fluent in F (e.g., both $B_1 p$ and $\neg B_1 p$ will become fluents), and the operators will describe how the mental model of our root agent should be updated. Formally, we define the classical encoding of a RP-MEP problem as follows:

Definition 2. Classical Encoding of RP-MEP

Let $\mathcal{B}_i$ and $\mathcal{N}_i$ be functions that map i's positive (resp. negative) belief from a set of RMLs KB to the respective fluents:

$$\mathcal{B}_i(KB) = \{l_\phi \mid B_i\phi \in KB\}$$
$$\mathcal{N}_i(KB) = \{l_\phi \mid \neg B_i\phi \in KB\}$$

Given a RP-MEP problem, $\langle \mathcal{P}, \mathcal{A}, Ag, \mathcal{I}, \mathcal{G}\rangle$ and a bound d on the depth of nested belief we wish to consider, we define the classical encoding as the tuple $\langle F, I, G, O\rangle$ such that:

$$F \overset{\text{def}}{=} \{l_\phi \mid \phi \in \mathcal{L}^{Ag,d}_{RML}\} \qquad I \overset{\text{def}}{=} \mathcal{B}_\star(\mathcal{I}) \qquad G \overset{\text{def}}{=} \mathcal{B}_\star(\mathcal{G})$$

and for every action $\langle \pi, \mathit{effects}\rangle$ in $\mathcal{A}$, we have a corresponding operator $\langle Pre_o, \mathit{eff}^+_o, \mathit{eff}^-_o\rangle$ in O such that:

$$Pre_o \overset{\text{def}}{=} \mathcal{B}_\star(\pi)$$
$$\mathit{eff}^+_o \overset{\text{def}}{=} \{(\langle \mathcal{B}_\star(\gamma_i), \mathcal{N}_\star(\gamma_i)\rangle \to l_\phi) \mid (\gamma_i, B_\star\phi) \in \mathit{effects}\}$$
$$\mathit{eff}^-_o \overset{\text{def}}{=} \{(\langle \mathcal{B}_\star(\gamma_i), \mathcal{N}_\star(\gamma_i)\rangle \to l_\phi) \mid (\gamma_i, \neg B_\star\phi) \in \mathit{effects}\}$$

3.2 Maintaining the Deductive Closure

Because of the direct correspondence, we will use the RML notation and terminology for the fluent atoms in F. The encoding, thus far, is a straight-forward adaptation of the RP-MEP definition that hinges on two properties: (1) there is a finite bound on the depth of nested belief; and (2) we restrict ourselves to representing RMLs and not arbitrary formulae. Crucially, however, we wish to maintain the assumption that the agents are internally consistent with respect to $KD45_n$. To accomplish this, we define a closure procedure, *Cl*, that deduces a new set of RMLs from an existing one under $KD45_n$:

Definition 3. RML Closure

Given an RML l, we define $Cl(l)$ to be the set of $KD45_n$ logical consequences of l computed as follows:

1. Rewrite l into *negation normal form* (NNF) (Bienvenu 2009), which is the equivalent formula in which negation appears only in front of propositional variables. For this we introduce an operator F s.t. $F_i\phi \equiv \neg B_i \neg\phi$.
2. Repeatedly apply the D axiom ($B_i\psi \supset F_i\psi$) to the NNF, resulting in the set of all RMLs that follow logically from ϕ using the D axiom. This can be done by simply replacing all combinations of occurrences of B_i with F_i; e.g., for the RML $B_iF_jB_kp$, the resulting set would be $\{F_iF_jB_kp,\ B_iF_jF_kp,\ F_iF_jF_kp\}$.
3. Invert the NNF by replacing all instances of $F_i\psi$ with $\neg B_i\neg\psi$ and eliminating double negation.

Note that to calculate the closure, we do not apply the positive (4) or negative (5) introspection actions of $KD45_n$, due to the equivalences in $KD45_n$ mentioned in Section 2. Proving the completeness of our RML closure is beyond the scope of this paper, but the soundness follows directly from the K and D axioms of $KD45_n$, and from the sound NNF re-writing rule (Bienvenu 2009).

Along with the requirement that an agent should never believe an RML and its negation, we have the following state constraints for the encoded planning problem:

$$\phi \in s \Rightarrow \neg\phi \notin s$$
$$\phi \in s \Rightarrow \forall\psi \in Cl(\phi), \psi \in s$$

The enforcement of such state constraints can either be achieved procedurally within the planner, or representationally. We choose the latter, appealing to a solution to the well-known ramification problem (e.g., (Pinto 1999; Lin and Reiter 1994)), representing these state constraints as *ancillary conditional effects* of actions that enforce the state constraints. The correctness of the resulting encoding is predicated on the assumption that the domain modeller provided a consistent problem formulation. The ancillary conditional effects for operator o are as follows:

$$(C \to l) \in \mathit{eff}_o^+ \Rightarrow (C \to \neg l) \in \mathit{eff}_o^- \tag{1}$$
$$(C \to l) \in \mathit{eff}_o^+ \Rightarrow \forall l' \in Cl(l),\ \ (C \to l') \in \mathit{eff}_o^+ \tag{2}$$

Example 2. Returning to our example, consider the effect of agent 1 telling secret s_1 to agent 2. Assuming there is no positive or negative condition for this effect to fire, the effect would be $(\langle\emptyset,\emptyset\rangle \to B_2s_1) \in \mathit{eff}^+$. Using (1) would create $(\langle\emptyset,\emptyset\rangle \to \neg B_2s_1) \in \mathit{eff}^-$ and (2) would create $(\langle\emptyset,\emptyset\rangle \to \neg B_2\neg s_1) \in \mathit{eff}^+$. Subsequently, (1) would fire again creating $(\langle\emptyset,\emptyset\rangle \to B_2\neg s_1) \in \mathit{eff}^-$. We can see already, with this simple example, that effects may cascade to create new ones.

3.3 Uncertain Firing and Removing Beliefs

To complete the faithful transformation of a RP-MEP problem to a classical problem, we must also consider the axioms that hold when updating the state due to the occurrence of an action. In particular, the following two issues remain: (1) the frame problem is solved only partially when we use the procedure listed above for updating a state in the encoded domain, and (2) removing belief about an RML may invalidate the state from being closed under $KD45_n$ (e.g., removing $\neg B_i\neg p$ while B_ip currently holds).

For the first issue, we appeal to a common technique in planning under uncertainty (e.g., (Petrick and Levesque 2002; Palacios and Geffner 2009)): when the conditions of a positive conditional effect are not believed to be false, the negation of the effect's result can no longer be believed. Intuitively, if an agent is unsure whether a conditional effect fires then it must consider the condition's effect possible, and thus no longer believe the negation of the effect. We create the following additional conditional effects for operator o:

$$\begin{aligned}&(\langle C^+, C^-\rangle \to l) \in \mathit{eff}_o^+ \Rightarrow \\ &\qquad (\langle\emptyset, \{\neg\phi \mid \phi \in C^+\} \cup C^-\rangle \to \neg l) \in \mathit{eff}_o^-\end{aligned} \tag{3}$$

The second issue is to ensure the state remains closed under $KD45_n$. If we remove an RML l, we should also remove any RML that could be used to deduce l. To compute the set of such RMLs, we use the contrapositive: $\neg l'$ will deduce l if and only if $\neg l$ deduces l' (i.e., $l' \in Cl(\neg l)$). We thus have the following additional conditional effects for operator o:

$$(C \to l) \in \mathit{eff}_o^- \Rightarrow \forall l' \in Cl(\neg l),\ (C \to \neg l') \in \mathit{eff}_o^- \tag{4}$$

Example 3. Consider a conditional effect for the action of agent 1 sharing their secret that stipulates if we, the root agent, think agent 1 is trustworthy (denoted as t_1), then we would believe agent 1's secret: $(\langle\{t_1\},\emptyset\rangle \to s_1) \in \mathit{eff}^+$. Using (3), we would derive the new negative effect $(\langle\emptyset,\{\neg t_1\}\rangle \to \neg s_1) \in \mathit{eff}^-$. Intuitively, if we are unsure about agent 1's trustworthiness, then we are unsure about their secret being false. On the other hand, consider the effect of an action informing us that we should no longer believe that agent 1 does not believe agent 2's secret: $(\langle\emptyset,\emptyset\rangle \to \neg B_1s_2) \in \mathit{eff}^-$. Using (4), we would have the additional effect $(\langle\emptyset,\emptyset\rangle \to B_1\neg s_2) \in \mathit{eff}^-$. If $B_1\neg s_2$ remained in our knowledge base, then so should $\neg B_1s_2$ assuming that our knowledge base is deductively closed.

With these extra conditional effects, we have a faithful encoding of the original RP-MEP problem.

Theorem 1. Our encoding is sound and complete with respect to RP-MEP. That is, a plan $\vec{o}$ will be found for a goal G from initial state I using our encoding if and only if $M, w \models \mathcal{I}$ implies $M, w \models [\vec{a}]\mathcal{G}$ for any (M, w), where M satisfies $KD45_n$ and $\vec{a}$ is the action sequence corresponding to $\vec{o}$.

Space precludes a formal proof, but this result follows from (1) the correctness of the newly added conditional effects, (2) the soundness and completeness of the *Cl* procedure, and (3) the use of a solution to the ramification problem to compile the properties into conditional effects.

4 Conditioned Mutual Awareness

Our specification of a RP-MEP problem and the subsequent encoding into classical planning allow us to specify a rich set of actions. Unlike traditional approaches that compile purely ontic action theories into ones that deal with belief (e.g., the work on conformant planning by Palacios and Geffner (2009)), we allow for arbitrary conditional effects that include nested belief both as conditions and as effects.

While expressive, manually encoding effects with nested belief can be involved due to the cascading of ancillary conditional effects. Here, we extend the scope of ancillary conditional effects to safely capture a common phenomenon in planning with nested belief: that of agents being mutually aware of the effects of actions.

Example 4. In our running example, if an agent enters a room, then we realize this as an effect: e.g., $(\langle\emptyset,\emptyset\rangle \rightarrow at\_1\_loc1) \in \mathit{eff}^+$. In many applications, other agents may also be aware of this: e.g., $(\langle\emptyset,\emptyset\rangle \rightarrow B_2 at\_1\_loc1) \in \mathit{eff}^+$. Perhaps we wish to predicate this effect on the second agent believing that it is also in this room: e.g., $(\langle\{B_2 at\_2\_loc1\},\emptyset\rangle \rightarrow B_2 at\_1\_loc1) \in \mathit{eff}^+$. It is this kind of behaviour of conditioned mutual awareness that we would like to capture in a controlled but automated manner.

By appealing to ancillary conditional effects, we will create new effects from existing ones. We have already demonstrated the ancillary conditional effects required for a faithful encoding to adhere to the axioms and state constraints we expect from our agent. We extend this idea here to capture the appealing property of *conditioned mutual awareness*. For simplicity, we describe conditioned mutual awareness in terms of the encoded problem, but assume the conditions for mutual awareness are optionally provided with a RP-MEP problem.

Definition 4. Condition for Awareness

We define $\mu_i^o \in F$ to be the condition for agent i to be aware of the effects of operator o. If need be, μ_i^o may be a unique fluent that is either always believed or never believed.

Intuitively, we want to assume that agent i is aware of every conditional effect of o only when agent i believes μ_i^o.

For a given set of fluents T, we define the shorthand $B_i T = \{B_i l \mid l \in T\}$ and $\neg B_i T = \{\neg B_i l \mid l \in T\}$ and model conditioned mutual awareness through the following two encoding rules for every agent $i \in Ag$ to derive new conditional effects:

$$(\langle C^+, C^-\rangle \rightarrow l) \in \mathit{eff}_o^+ \Rightarrow$$
$$(\langle B_i C^+ \cup \neg B_i C^- \cup \{B_i\, \mu_i^o\}, \emptyset\rangle \rightarrow B_i l) \in \mathit{eff}_o^+ \tag{5}$$
$$(\langle C^+, C^-\rangle \rightarrow l) \in \mathit{eff}_o^- \Rightarrow$$
$$(\langle B_i C^+ \cup \neg B_i C^- \cup \{B_i\, \mu_i^o\}, \emptyset\rangle \rightarrow \neg B_i l) \in \mathit{eff}_o^+ \tag{6}$$

Note that each form of ancillary conditional effect adds a new positive conditional effect. In the positive case, we believe that the agent i has a new belief $B_i l$ if we believe that agent i had the prerequisite belief for the effect to fire. In the negative case, we would believe that the agent no longer holds the belief, but because we take a perspectival view, it is encoded as a positive conditional effect – i.e., we would believe $\neg B_i l$. For instance, the ancillary conditional effect from our working example says that we should no longer believe the negation of agent 1's secret if we do not believe agent 1 is untrustworthy (see Example 3), would create the following ancillary conditional effect:

$$(\langle\emptyset, \{\neg t_1\}\rangle \rightarrow \neg s_1) \in \mathit{eff}^- \Rightarrow$$
$$(\langle\{\neg B_2 \neg t_1\}, \emptyset\rangle \rightarrow \neg B_2 \neg s_1) \in \mathit{eff}^+.$$

We restrict the application of the above rules by applying them only if the following two conditions are met: (1) every RML in the newly created effect has a nested depth smaller than our bound d; and (2) if we are applying the above rule for agent i to a conditional effect $(C \rightarrow l) \in \mathit{eff}_o^-$, then $l \notin \{B_i l', \neg B_i l'\}$. The first restriction bounds the number of conditional effects while the second prevents unwanted outcomes from introspection. To see why this exception is required, consider the example of a pair of conditional effects for an action where we discover agent 1 may or may not believe s_2 (i.e., we should forget any belief about what agent 1 believes regarding s_2). Omitting μ_1^o for clarity, we have the following *negative* conditional effects:

$$(\langle\emptyset,\emptyset\rangle \rightarrow \neg B_1 s_2) \quad (\langle\emptyset,\emptyset\rangle \rightarrow B_1 s_2)$$

If we were to apply the above rules with agent 1, we would add two *positive* ancillary conditional effects:

$$(\langle\emptyset,\emptyset\rangle \rightarrow \neg B_1 \neg B_1 s_2) \quad (\langle\emptyset,\emptyset\rangle \rightarrow \neg B_1 B_1 s_2)$$

which subsequently would simplify to the following conditional effects (given that we combine successive modalities of the same agent index under $\mathrm{KD45}_n$):

$$(\langle\emptyset,\emptyset\rangle \rightarrow B_1 s_2) \quad (\langle\emptyset,\emptyset\rangle \rightarrow \neg B_1 s_2)$$

Thus, the resulting effects would indicate that the agent reaches an inconsistency with its own belief. To avoid this issue, we apply rule (6) only when the effect is not a belief (negative or positive) of the corresponding agent.

Because we can assume that the specification of conditioned mutual awareness is given and computed in the original RP-MEP specification, Theorem 1 continues to hold.

5 Projection to Reason As Others

It is natural that an agent may want to reason about what another believes or may come to believe, allowing queries such as, "Does Sue believe that Bob believes that Sue believes a plan exists?". We construe the term *virtual agent* to be the list of agents we wish to have the root agent reason as. Here, [Bob, Sue] is the virtual agent assuming that Sue is the root agent: i.e., we want Sue to reason as if she was Bob reasoning as if he was Sue.

To reason as a virtual agent, we require two items: (1) the assumed mental model of the virtual agent's initial state, and (2) the virtual agent's view of the operators. We assume that the goal, set of agents, and operator preconditions remain the same. To create the new initial state for the virtual agent, we use projection:

Definition 5. Agent Projection

Given a state s, the *agent projection* of s with respect to a vector of agents $\vec{Ag}$, denoted as $Proj(s, \vec{Ag})$, is defined as:

$$\begin{cases} \{\phi \mid B_i\phi \in s\} & \text{if } \vec{Ag} = [i] \\ Proj(Proj(s,[i]), \vec{Ag}') & \text{if } \vec{Ag} = [i] + \vec{Ag}' \end{cases}$$

Essentially, agent projection repeatedly filters the set of RMLs according to the appropriate agent and strips the belief modality from the front of the RML. When projecting a planning problem, we project the initial state using agent projection – giving us the believed mental state of the virtual agent – and additionally project the effects of every operator. Because we allow for heterogeneous agents with respect to their view on operator effects, we first must decide which conditional effects to keep for the projection. For a particular agent i, these are the effects *uniform in i*.

Definition 6. Uniform Conditional Effect

We say that an RML is uniform in i if the RML is a condition for awareness or it begins with either the modality B_i or $\neg B_i$. A set of fluents is uniform in i iff every RML in the set is uniform in i. Finally, the set of conditional effects uniform in i for operator o are all those $(\langle C^+, C^-\rangle \rightarrow l) \in \mathit{eff}_o^+$ such that C^+ is uniform in i, l is uniform in i, and $C^- = \emptyset$.

The projection of an operator for agent i will retain all those conditional effects that are uniform in i. Note that this discards all negative conditional effects. Once we have the set of uniform conditional effects, we project each effect $(\langle C^+, \emptyset\rangle \rightarrow l)$ in the set for the agent i to be defined as:

$$(\ \langle \{\phi \mid B_i\phi \in C^+\},\ \{\phi \mid \neg B_i\phi \in C^+\} \rangle \rightarrow l'),$$

where $l = B_i l'$ and the projected effect is in eff_o^+ or $l = \neg B_i l'$ and the projected effect is in eff_o^-.

The intuition behind our definition of conditional effects uniform for agent i, is that we consider only those effects that we (the current agent) believe agent i will reason with. If a conditional effect has a negative condition (i.e., C^- is non-empty), then that is a condition that involves our own lack of belief and not the lack of belief for agent i (the latter would exist as an RML starting with $\neg B_i$ in C^+). Similarly, negative conditional effects describe how *we* remove belief, and not how agent i would update their belief. Paired with the ability to add conditioned mutual awareness, the projection of effects for a particular agent can target precisely those effects we want to keep.

We generate a new initial state for a particular virtual agent using the agent projection procedure, and we generate a new set of operators by repeatedly applying the above procedure for operator projection. Additionally, because we assume that the nestings of modalities of the same agent index are combined (cf. Section 2), we process the operator preconditions and goal slightly to accommodate for the new root agent perspective assumed for the final agent i in the virtual agent list: B_i is removed from the start of precondition and goal RMLs while any RML of the form $\neg B_i\phi$ is converted to a negative precondition or goal RML ϕ. The strength in projecting away effects is that it can simplify the domain greatly – any conditional effect not uniform in the projected agent will be pruned from the domain prior to planning. Combining ancillary conditional effects and projection allows us to answer a complex suite of queries for the nested belief of agents.

6 Preliminary Evaluations

We implemented the scheme above to convert a RP-MEP planning problem into a classical planning problem, which can be subsequently solved by any planner capable of handling negative preconditions and conditional effects. The compiler consumes a custom format for the RP-MEP problems and can either simulate the execution of a given action sequence or call the Fast Downward planner (Helmert 2006).

We have verified the model of the pre-existing *Thief* problem, and all of the existing queries considered in the previous literature posed to demonstrate the need for nested reasoning (e.g., those found in (Löwe, Pacuit, and Witzel 2011)) are trivially solved in a fraction of a second. As a more challenging test-bed, we modelled a setting that combines the *Corridor* problem (Kominis and Geffner 2014) and the classic *Gossip* problem (Entringer and Slater 1979). In the new problem, *Grapevine*, there are two rooms with all agents starting in the first room. Every agent believes their own secret to begin with, and the agents can either move between rooms or broadcast a secret they believe. Movement is always observed by all, but through the use of conditioned mutual awareness the sharing of a secret is only observed by those in the same room. This problem allows us to pose a variety of interesting goals ranging from private communication (similar to the Corridor problem) to goals of misconception in the agent's belief (e.g., $G = \{B_a s_b, B_b \neg B_a s_b\}$).

As a preliminary investigation, we varied some of the discussed parameters and report on the results in Table 1 (the first Corridor problem corresponds to the one presented by Kominis and Geffner). The largest bottleneck stems from the depth of nested knowledge, as the number of newly introduced fluents is exponential in d. The planning process is typically fast, and moving forward we hope to reduce the compilation time by only generating fluents and conditional effects that are relevant to achieving the goal.

7 Related Work

There is a variety of research related to the ideas we have presented, and we cover only the most closely related here. Research into DEL (van Ditmarsch, van der Hoek, and Kooi 2007), and more recently DEL planning (e.g., (Bolander and Andersen 2011)), deals with how to reason about knowledge or belief in a setting with multiple agents. Focus in this area is primarily on the logical foundation for updating an epistemic state of the world according to physical (ontic) and non-physical (epistemic) actions, as well as identifying the classes of restricted reasoning that are tractable from a theoretical standpoint (Löwe, Pacuit, and Witzel 2011). While DEL techniques are more expressive than our approach in terms of the logical reasoning that an agent can achieve in

Problem	$\lvert Ag \rvert$	d	$\lvert F \rvert$	$\lvert \vec{o} \rvert$	Time (s) Plan	Time (s) Total
Corridor	3	1	70	5	0.01	1.14
	7	1	150	5	0.01	1.32
	3	3	2590	5	0.01	38.80
Grapevine	4	1	216	3	0.01	1.26
	3	2	774	4	0.01	3.44
	4	2	1752	7	0.16	10.81

Table 1: Results for various Corridor and Grapevine problems. Ag, d, F, and $\vec{o}$ are as above. Plan time is the time spent solving the encoded problem, while Total time additionally includes the encoding and parsing phases.

theory, this expressiveness increases the computational complexity of the reasoning. In particular, the practical synthesis of DEL plans remains an unsolved problem.

On the other end of the spectrum, a range of techniques exist for planning with partial observability (e.g., (Brafman and Shani 2012; Bonet and Geffner 2014)). These techniques typically represent the individual knowledge of facts about the world (as opposed to belief): the agent can "know p holds" (i.e., Kp), "know p does not hold" (i.e., $K\neg p$), or "not know the value of p" (i.e., $\neg Kp \wedge \neg K \neg p$), but do not extend to the multi-agent case. The use of a knowledge modality was extended to be predicated on assumptions about the initial state, leading to effective techniques for conformant and contingent planning (Palacios and Geffner 2009; Albore, Palacios, and Geffner 2009).

Others have investigated restricted classes of syntactic knowledge bases for tractable reasoning, including the 0-approximation semantics (Baral and Son 1997), and the notion of *Proper Epistemic Knowledge Bases* (PEKBs) (Lakemeyer and Lespérance 2012). Like ours, these works restrict the type of knowledge that is stored about a single agent, such as not permitting disjunctive knowledge (e.g., the agent "knows either p or q holds"). PEKBs extend this to the multi-agent setting, and in a sense the preconditions, goals, and states of our work can be viewed as PEKBs.

The work most related to ours is that of Kominis and Geffner (2014). They share the same general motivation of bridging the rich fields of epistemic reasoning and automated planning by using classical planning over multi-agent epistemic states. However, the two approaches are fundamentally different and as a result each comes with its own strengths and shortcomings. The largest difference is our choice to focus on belief rather than knowledge – for us, modelling the possibility of incorrect belief is essential. Kominis and Geffner assume that all agents start with common initial knowledge, and most strongly that all action effects are commonly known to all agents (while we can model this, it is not required). Conversely, they are able to handle arbitrary formulae, including disjunctive knowledge, while we are restricted to reasoning with RMLs. Moving forward, we hope to explore how we can combine ideas from both approaches.

8 Concluding Remarks

We have presented a model of planning with nested belief, and demonstrated how a syntactically restricted subclass of this expressive problem can be compiled into a classical planning problem. Despite the restricted form, we are able to model complex phenomena such as public or private communication, commonly observed action effects, and non-homogeneous agents (each with their own view of how the world changes). Our focus on belief (as opposed to knowledge) provides a realistic framework for an agent to reason about a changing environment where knowledge cannot be presumed. We have additionally demonstrated how to pose queries *as if* we were other agents, taking our belief of the other agents into account. To solve this expressive class of problems, we appeal to existing techniques for dealing with ramifications, and compile the problem into a form that classical planning can handle.

In future work, we hope to expand the work in two key directions. First, we would like to explore other forms of ancillary conditional effects similar to the conditioned mutual awareness to give the designer greater flexibility during modelling (e.g., with concepts such as teamwork protocols or social realities). Second, we want to formalize the connection between general multi-agent epistemic planning and the syntactic restriction that we focus on encoding. We hope to provide an automated sound (but incomplete) approximation of an arbitrary MEP problem into a RP-MEP problem.

Acknowledgements This research is partially funded by Australian Research Council Discovery Grant DP130102825, *Foundations of Human-Agent Collaboration: Situation-Relevant Information Sharing*, and by the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

Albore, A.; Palacios, H.; and Geffner, H. 2009. A translation-based approach to contingent planning. In *IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, Pasadena, California, USA, July 11-17, 2009*, 1623–1628.

Aucher, G., and Bolander, T. 2013. Undecidability in epistemic planning. In *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*.

Baral, C., and Son, T. C. 1997. Approximate reasoning about actions in presence of sensing and incomplete information. In *Logic Programming, Proceedngs of the 1997 International Symposium, Port Jefferson, Long Island, NY, USA, October 13-16, 1997*, 387–401.

Bienvenu, M. 2009. Prime implicates and prime implicants: From propositional to modal logic. *Journal of Artificial Intelligence Research* 36(1):71–128.

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.

Bonet, B., and Geffner, H. 2014. Belief tracking for planning with sensing: Width, complexity and approximations.

Journal of Artificial Intelligence Research (JAIR) 50:923–970.

Brafman, R. I., and Shani, G. 2012. Replanning in domains with partial information and sensing actions. *Journal of Artificial Intelligence Research (JAIR)* 45:565–600.

Brenner, M., and Nebel, B. 2009. Continual planning and acting in dynamic multiagent environments. *Autonomous Agents and Multi-Agent Systems* 19(3):297–331.

Entringer, R. C., and Slater, P. J. 1979. Gossips and telegraphs. *Journal of the Franklin Institute* 307(6):353–360.

Fagin, R.; Halpern, J. Y.; Moses, Y.; and Vardi, M. Y. 1995. *Reasoning about knowledge*, volume 4. MIT press Cambridge.

Ghallab, M.; Nau, D.; and Traverso, P. 2004. *Automated planning: theory & practice*. Elsevier.

Helmert, M. 2006. The fast downward planning system. *Journal of Artificial Intelligence Research (JAIR)* 26:191–246.

Kominis, F., and Geffner, H. 2014. Beliefs in multiagent planning: From one agent to many. In *Workshop on Distributed and Multi-Agent Planning*, 62–68.

Lakemeyer, G., and Lespérance, Y. 2012. Efficient reasoning in multiagent epistemic logics. In *ECAI 2012 - 20th European Conference on Artificial Intelligence. Including Prestigious Applications of Artificial Intelligence (PAIS-2012) System Demonstrations Track, Montpellier, France, August 27-31 , 2012*, 498–503.

Lin, F., and Reiter, R. 1994. State constraints revisited. *J. Log. Comput.* 4(5):655–678.

Liu, Y., and Wen, X. 2011. On the progression of knowledge in the situation calculus. In *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence, Barcelona, Catalonia, Spain, July 16-22, 2011*, 976–982.

Löwe, B.; Pacuit, E.; and Witzel, A. 2011. DEL planning and some tractable cases. In *Logic, Rationality, and Interaction - Third International Workshop, LORI 2011, Guangzhou, China, October 10-13, 2011. Proceedings*, 179–192.

Palacios, H., and Geffner, H. 2009. Compiling uncertainty away in conformant planning problems with bounded width. *Journal of Artificial Intelligence Research (JAIR)* 35:623–675.

Petrick, R. P. A., and Levesque, H. J. 2002. Knowledge equivalence in combined action theories. In *Proceedings of the Eights International Conference on Principles and Knowledge Representation and Reasoning (KR-02), Toulouse, France, April 22-25, 2002*, 303–314.

Pinto, J. 1999. Compiling ramification constraints into effect axioms. *Computational Intelligence* 15:280–307.

Reiter, R. 2001. *Knowledge in action: logical foundations for specifying and implementing dynamical systems*, volume 16. MIT press Cambridge.

van Ditmarsch, H.; van der Hoek, W.; and Kooi, B. P. 2007. *Dynamic epistemic logic*, volume 337. Springer.

van Eijck, J. 2004. Dynamic epistemic modelling. *Manuscript, CWI, Amsterdam*.