Epistemic Plan Recognition

Maayan Shvo*
Department of Computer Science,
University of Toronto
Toronto, Canada
maayanshvo@cs.toronto.edu

Toryn Q. Klassen*
Department of Computer Science,
University of Toronto
Toronto, Canada
toryn@cs.toronto.edu

Shirin Sohrabi
IBM Research
Yorktown Heights, United States
ssohrab@us.ibm.com

Sheila A. McIlraith*
Department of Computer Science,
University of Toronto
Toronto, Canada
sheila@cs.toronto.edu

ABSTRACT

The plan recognition task is to infer an actor's plan and goal given observations about its behavior. We submit that in some cases, for plan recognition to be effective and complete, it must appeal to a notion of epistemics to i) recognize epistemic goals, where the actor is trying to achieve some state of knowledge or belief; and ii) model the observer, and its knowledge of the actor, as first class elements of the plan recognition process. To this end, we formalize the notion of *Epistemic Plan Recognition*, which builds on two growing areas of research: epistemic planning and plan recognition. Our epistemic plan recognition specification appeals to an epistemic logic framework to represent agent beliefs. To realize our specification, we cast the epistemic plan recognition problem as an epistemic planning problem, whose solutions can be generated using existing epistemic planning tools. Finally, we evaluate our approach by utilizing and comparing existing epistemic planners on a diverse set of epistemic plan recognition problems.

KEYWORDS

Plan Recognition; Epistemic Planning

ACM Reference Format:
Maayan Shvo, Toryn Q. Klassen, Shirin Sohrabi, and Sheila A. McIlraith. 2020. Epistemic Plan Recognition. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 9 pages.

1 INTRODUCTION

Plan Recognition (PR) - the task of inferring the plan and goal of an actor based on observations - can be seen as an exercise in the observer's Theory of Mind (i.e., an agent's ability to attribute mental states to itself and to others [40]). That is, in order to explain the observed behavior of the actor, the observer attributes to the latter various mental states - beliefs, plans, and goals. The recognition process is thus inherently *epistemic* in that it is determined by the beliefs of the observer about the beliefs, plans, and goals of the actor. We therefore take the view (espoused by Pollack [39] and others) that plan recognition can necessitate representing and reasoning about the potentially distinct beliefs of the observer and the actor, and in particular that the observer may be required to assume the perspective of the actor in order to effectively recognize what the actor is doing. We further observe that the goals being pursued by the actor may be *ontic*—related to changing the state of the world, *epistemic*—related to changing its own (or another agent's) state of belief or knowledge, or perhaps both.

To advance this view of plan recognition, we introduce and formalize the notion of *Epistemic Plan Recognition* (EPR) in which we explicitly represent and reason about agents' (nested) beliefs about the world, and the effects of actions on the world and on the beliefs of other agents. To realize EPR, we appeal to the burgeoning field of *Epistemic Planning* which exploits state-of-the-art AI planning techniques to generate plans to achieve epistemic goals (e.g., [5, 11, 18, 21, 24, 33, 38]). By appealing to epistemic logic we are able to: (1) model the observer and its knowledge of the actor's beliefs and capabilities as first class elements of the plan recognition process, and (2) recognize plans with epistemic as well as ontic goals, or their combination. Modelling the observer is important as it enables reasoning about its own, and the actor's, beliefs, ignorance, and misconceptions relating to its environment and to the beliefs and capabilities of other agents.

Plan recognition, as a field of research, has a long history and was originally seen as an intersection of psychology and AI [44]. Early accounts of PR largely utilized (possibly hierarchical) plan libraries in order to best match a sequence of observations to a particular plan (e.g., [19]). More recent work in the field has dispensed with the need for a plan library by casting the PR problem as a planning problem and leveraging advances in AI planning research (e.g., [42, 47]). However, most previous work did not appeal to a notion of epistemics to allow for the recognition of epistemic goals and to explicitly model the observer. While Talamadupula et al. [49] model aspects of agent beliefs in their work, they do not address the recognition of epistemic goals nor do they utilize epistemic planning tools. PR has also been studied within the vast body of work on Belief-Desire-Intention (BDI) agents and architectures, where agent beliefs are explicitly modelled (e.g., [46]). However, BDI approaches have typically required agent plans to be specified in advance. Pre-defined libraries are especially prohibitive when

*The author is also affiliated with the Vector Institute in Toronto, Canada.

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), , May 9–13, 2020, Auckland, New Zealand.

agents have misconceptions and generate invalid plans [39]. Instead, enabled by the advent of epistemic planning research, we appeal to the flexibility of generative epistemic planning techniques to generate plans.

The main contributions of this paper are: (1) a formalization of EPR which adds an important dimension to the recognition process by appealing to a notion of epistemics to allow for the recognition of epistemic goals and to model the observer and its knowledge of the actor as first class elements of the recognition process; (2) a planner-independent computational realization of EPR as epistemic planning that synthesizes formalism and computational techniques from both epistemic planning and PR and enables the use of existing planning tools; and (3) an evaluation of our approach on a set of EPR problems from a diversity of domains which offers a comparison between 3 state-of-the-art epistemic planners.

2 PRELIMINARIES

In this section, we provide epistemic logic background, define the Multi-agent Epistemic Planning (MEP) problem, and describe some of the approaches to modelling it.

KD45$_n$. We first present the multi-agent modal logic KD45$_n$ [13] which we appeal to in our specification of EPR. Let Ag and $\mathcal{P}$ be finite sets of agents and atoms, respectively. We use ϕ and ψ to represent formulae and $\top$ and $\bot$ to represent *true* and *false*, respectively. The language $\mathcal{L}$ of multi-agent modal logic is generated by the following BNF:

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

where $p \in \mathcal{P}$, $i \in Ag$, $\phi, \psi \in \mathcal{L}$ and $B_i\phi$ should be interpreted as "*agent i believes* ϕ." We choose to represent the belief (rather than knowledge) modality here so that we can model the false beliefs of agents. The semantics for formulae in $\mathcal{L}$ is given by Kripke models [13] which are triplets, $M = \langle W, R, V\rangle$, containing a set of worlds, accessibility relations between the worlds for each of the agents ($R = \{R_i \mid i \in Ag\}$), and a valuation map, $V : W \to 2^P$. When an agent i is at world $w \in W$, M determines, given the accessibility relations in R_i pertaining to w, what worlds the agent considers possible. A formula ϕ is true in a world w of a Kripke model $M = \langle W, R, V\rangle$, written $M, w \models \phi$, under these, inductively-defined conditions:

- $M, w \models p$ for an atom p, *iff* $p \in V(w)$,
- $M, w \models \neg\phi$, *iff* $M, w \not\models \phi$,
- $M, w \models \phi \wedge \psi$, *iff* both $M, w \models \phi$ and $M, w \models \psi$,
- $M, w \models B_i\phi$, *iff* $M, w' \models \phi \quad \forall w' \in W$ s.t. $R_i(w, w')$

We say that ϕ is satisfiable if there is a Kripke model M and a world w of M s.t. $M, w \models \phi$. Further, we say that ϕ entails ψ, written $\phi \models \psi$, if for any Kripke model M, $M, w \models \phi$ entails $M, w \models \psi$. Next, we assume some constraints on the Kripke model, with particular properties of belief, as discussed in Fagin et al. [13]. Namely, we assume that the Kripke model is *serial* ($\forall w\, \exists v\, R(w, v)$), *transitive* ($R(w, v) \wedge R(v, u) \Rightarrow R(w, u)$) and *Euclidean* ($R(w, v) \wedge R(w, u) \Rightarrow R(v, u)$), with the resulting properties of belief: i. $B_i\phi \wedge B_i(\phi \Rightarrow \psi) \Rightarrow B_i\psi$ (K - Distribution); ii. $B_i\phi \Rightarrow \neg B_i\neg\phi$ (D - Consistency); iii. $B_i\phi \Rightarrow B_iB_i\phi$ (4 - Positive Introspection); and iv. $\neg B_i\phi \Rightarrow B_i\neg B_i\phi$ (5 - Negative Introspection). These axioms, together, form the KD45$_n$ system where n signifies multiple agents in the environment.

Epistemic Planning. We appeal to epistemic planning both to specify and to realize EPR. As discussed by Bolander [4], approaches to epistemic planning can be characterized as either *syntactic* (e.g., [18, 33]) or *semantic* (e.g., [5, 24]). The syntactic approach represents states as knowledge bases which are sets of logical formulae, while the semantic approach represents states as semantical objects, and manipulates them directly. To simplify the exposition, we choose to focus here on a syntactic approach and, consequently, the initial knowledge base (KB) is an arbitrary epistemic logic formula, rather than a Dynamic Epistemic Logic (DEL) epistemic state (e.g., [5]). We appeal to a *multi-agent* setting in order to represent the beliefs of the observer, the actor, and possibly of other agents as well.

Definition 2.1 (MEP Problem). A **Multi-agent Epistemic Planning problem** is a tuple $\langle Q, \mathcal{I}, G\rangle$ where $Q = \langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag\rangle$ is the MEP domain comprising sets of atoms $\mathcal{P}$, actions $\mathcal{A}$, action descriptions $\mathcal{D}$, and agents Ag, together with the problem instance description comprising the initial KB, $\mathcal{I} \in \mathcal{L}$, and the goal condition $G \in \mathcal{L}$, where $\mathcal{L}$ is the language of multi-agent modal logic.

Following Huang et al. [18] and Muise et al. [33], and since we are interested in a PR setting, we model the MEP problem from the perspective of an observer, Obs $\in Ag$. $\mathcal{I}$ and $\mathcal{D}$ therefore represent the observer's beliefs about the world and the actions in $\mathcal{A}$, including its beliefs about other agents' (in Ag) beliefs about the world and the actions. For simplicity, all formulae are assumed to be implicitly prefixed by B_{Obs}. $\mathcal{D}$ is a set of size $|Ag| \times |\mathcal{A}|$ where each element in $\mathcal{D}$ is an action description corresponding to an action $a \in \mathcal{A}$ and an agent $i \in Ag$. An action description for action $a \in \mathcal{A}$ is a tuple $\langle Pre, \{(\gamma_1, \epsilon_1), \ldots, (\gamma_k, \epsilon_k)\}\rangle$, where $Pre \in \mathcal{L}$ is called the precondition of a, $\gamma_i \in \mathcal{L}$ is the condition of a conditional effect, and $\epsilon_i \in \mathcal{L}$ is called the effect of a conditional effect. For example, the conditional effect (q, B_ip) for action a (corresponding to Obs $\in Ag$) denotes that the observer believes that agent i will believe p following the execution of action a, provided the observer believed q at the time a was executed. $\mathcal{A}$ is shared by all agents.

To change an agent's beliefs about the state of the world following the execution of an action, we appeal to a progression operator $prog_i(\phi, a)$ which progresses $\phi \in \mathcal{L}$ wrt an action $a \in \mathcal{A}$ in the context of agent $i \in Ag$. That is, ϕ is progressed wrt the action description corresponding to a and i in $\mathcal{D}$. We define the progression of a formula in the context of a particular agent, i, in order to be able to capture the change in agent i's beliefs following the execution of an action. This is important because agents can differ in how they conceive the effects and preconditions of actions and, as with MEP, we want our EPR system to be able to assume the perspective of different agents. For example, while the observer may believe that some action a has the effect p, it may also believe that agent i believes that a's effect is q. The progression operator is assumed to be sound and complete wrt to the chosen fragment of the logic and may be realized in various ways in the context of epistemic planning. For instance, Muise et al.'s epistemic planner RP-MEP [33] builds on Proper Epistemic Knowledge Bases (PEKBs) (Lakemeyer and Lespérance [23] building on Liu et al. [27] and Levesque [25]) with follow up work by Muise et al. [32]. Miller and Muise [31] define a belief update mechanism for PEKBs that they have used as the basis of a progression operator over PEKBs. Progression has also been defined by Huang et al. [18].

We use the shorthand $prog_i(\phi, [a_1, ..., a_n])$ or $prog_i(\phi, \pi)$ to denote the progression of ϕ wrt a sequence of actions $\pi = a_1, ..., a_n$ in the context of agent i. A plan π **achieves** a goal G wrt agent $i \in Ag$ *iff* $prog_i(\mathcal{I}_i, \pi) \models G$. $\mathcal{I}_i \in \mathcal{L}$ is the observer's beliefs about agent i's initial beliefs. A **solution** to an MEP problem is a sequence of actions π such that π achieves G wrt the observer. Here we assume that the *cost* of π equals the number of actions in the sequence.

3 EPISTEMIC PLAN RECOGNITION

The task of plan recognition is to infer a plan and goal that account for the observed behavior of an actor. As argued in Section 1, a general account of plan recognition should provide a means of modelling the beliefs of the observer. To this end, in defining the EPR problem we include the observer (Obs $\in Ag$) and the actor (Act $\in Ag$) in the set of agents, Ag. The goals are drawn from a set of possible goals, $\mathcal{G}$, the actor may be pursuing. Finally, the observed behavior comprises both actions and properties of the state of the system. In particular, a sequence of observations is a sequence of tuples $(\alpha_1, \phi_1), \ldots, (\alpha_m, \phi_m)$. Each α_i corresponds to the observation of an action, $a \in \mathcal{A}$, and each $\phi_i \in \mathcal{P}$ corresponds to the observation of some properties of state immediately following the execution of α_i. ϕ_i is a conjunction of literals drawn from $\mathcal{P}$. For example, the observation (*leave(John, room)*, *empty(room)*) signifies that the observer had observed that the room is empty after observing John leaving the room. In cases where only properties are observed, α_i is empty. In cases where an action is observed but no state properties, ϕ_i is $\top$, i.e., *true*.

Definition 3.1 (EPR Problem). An **Epistemic Plan Recognition problem** is a tuple $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O \rangle$, where $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag \rangle$ is an MEP domain, $\mathcal{I} \in \mathcal{L}$ is the initial KB, $\mathcal{G}$ is a set of possible goals $G \in \mathcal{L}$, and $O = o_1, ..., o_m$, is a sequence of observations. Each observation o_i is a pair (α_i, ϕ_i) comprising zero or one observed actions, $\alpha_i \in \mathcal{A}$, together with ϕ_i, a conjunction of literals drawn from $\mathcal{P}$, or $\top$.

Given an EPR problem, $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O \rangle$, a sequence of actions $\pi = a_1, \ldots, a_n$ **satisfies** observations $O = (\alpha_1, \phi_1), \ldots, (\alpha_m, \phi_m)$ if there is a monotonic function f mapping the observation indices $j = 1, ..., m$ into action indices $i = 1, ..., n$ such that $a_{f(j)} = \alpha_j$ (trivially satisfied when α_j is empty), and $prog_{\text{Obs}}(\mathcal{I}, [a_1, \ldots, a_{f(j)}]) \models \phi_j$ for $j = 1, \ldots, m$. Note that O is satisfied from the observer's perspective.

Definition 3.2 (EPR Solution). Given an EPR problem, $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O \rangle$, a **solution** is a pair (π, G), where $G \in \mathcal{G}$ is a goal and π is a sequence of actions – a plan – that satisfies O.

We later define possible relationships between the plan component π and the goal component G (which is unconstrained in Definition 3.2), characterized by the observer's beliefs about whether or not π achieves G from the perspectives of the actor and observer.

Example. Consider a scenario where Alice (denoted as Act $\in Ag$) has the goal of bringing a cake to her friend Eve's party and an observer (Obs $\in Ag$) is trying to infer her plan and goal. Alice is observed going to the store, buying a cake, and taking it to the party. We partially model this scenario as an EPR problem:

$Ag = \{\text{Obs}, \text{Act}\}$

$\mathcal{I} \models at(\text{Act}, \text{Home}) \land B_{\text{Act}}\, at(\text{Act}, \text{Home}) \land B_{\text{Act}} \neg at(\text{Cake}, \text{Party})$

$O =$ (*takeBus(*Act,Home,Store*)*, *at(*Act,Store*)*), (*buy(*Act,Cake*)*, *have(*Act,Cake*)*), (*take*(Act,Bag,Party), $\top$)

$\{at(\text{Cake},\text{Home}),\ at(\text{Cake},\text{Work}),\ at(\text{Cake},\text{Party})\} \subseteq \mathcal{G}$

$\mathcal{I}$ entails many things and we only list a small portion of those. A possible solution to this EPR problem is the pair (π, G) where the presumed plan π consists of the actions in O and the presumed goal G is *at*(Cake,Party). While the plan π in this case achieves the goal G both wrt the observer and wrt the actor, Pollack [39] identified that *"inferring another agent's plan means figuring out what actions she 'has in mind,' and she may well be wrong about the effects of those intended actions"*. We therefore distinguish between the actor's presumed plan π and the underlying intent of the plan, i.e., the presumed goal G. When the actor is wrong about the outcomes of her plan, such a plan may be *ill-formed* [39].

Definition 3.3. Given an EPR problem $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O \rangle$, we say that a plan π is **ill-formed** wrt a goal $G \in \mathcal{L}$ *iff* π achieves G wrt the actor but does not achieve G wrt the observer.

Example. Now consider that the observer knows that the party's address is Address1 and further knows that Alice *falsely* believes that the address is Address2 (i.e., $\mathcal{I} \models loc(\text{Party},\text{Address1}) \land B_{\text{Act}}\, loc(\text{Party},\text{Address2})$). Alice's presumed plan, π, is now *ill-formed* since following its execution Alice will believe that she is at the party (i.e., $B_{\text{Act}}\, at(\text{Act},\text{Party})$) since she believes *loc*(Party,Address2) while the observer will believe $\neg at(\text{Act},\text{Party})$ following π's execution since it believes *loc*(Party,Address1). While we focus in this paper on the recognition task, an assistive observer can reason that Alice's plan is ill-formed as soon as she is observed to be heading towards Address2 rather than Address1 and can help her 'correct course'. Pollack also discusses the notion of incoherent queries[1] and we extend the discussion to plans.

Definition 3.4. Given an EPR problem $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O \rangle$, we say that a plan π is **incoherent** wrt a goal $G \in \mathcal{L}$ *iff* π does not achieve the goal G both wrt the actor and wrt the observer.

An incoherent plan does not 'make sense' to the observer since it fails to achieve the presumed goal even from the actor's presumed perspective. For instance, a plan consisting of Alice disposing of the cake after buying it would likely be incoherent to the observer wrt *at*(Cake,Party).

Example. Now consider that Alice wants to surprise Eve and so does not want her to know that she has brought cake to the party. We include Eve as an agent in Ag so that our representational framework can represent such goals in $\mathcal{G}$ (i.e., $\neg B_{Eve}\, at(\text{Cake},\text{Party}) \land at(\text{Cake},\text{Party})$) and augment $\mathcal{G}$ appropriately. Alice initially believes that Eve is at the party (i.e., $\mathcal{I} \models B_{\text{Act}}\, at(\text{Eve},\text{Party})$) and is observed to *hide* the cake in a dark bag after buying it, such that following the *hide* action, Alice's beliefs include *hidden*(Cake) $\land \neg$*hidden*(Bag) $\land \neg$*hidden*(Act). Next, following the execution of *take*(Act,Bag,Party), Alice's beliefs include *at*(Cake,Party) $\land$ *at*(Act,Party) $\land$ *at*(Bag,Party) and since the bag and Alice are not

[1] In Pollack's work, the observer's task is to infer the actor's (possibly invalid) plan, underlying the latter's dialogue queries.

hidden, Alice believes that every agent at the party now believes *at*(Act,Party) $\wedge$ *at*(Bag,Party). However, Alice does not believe the agents at the party have changed their beliefs about *at*(Cake,Party) since she believes *hidden*(Cake). Since Alice believes that Eve is at the party (and that Eve initially does not believe *at*(Cake,Party)), she also believes $\neg B_{Eve}\, at$(Cake,Party) following the execution of her plan. Importantly, to recognize Alice's epistemic goal an observer must reason not only about Alice's beliefs about the world, but also about her nested beliefs pertaining to Eve's beliefs.

Adequacy. EPR, like any PR system, is limited by the veracity and completeness of the observer's knowledge and beliefs—the set of possible goals, the sequence of observations, and, importantly, the veracity and completeness of the observer's beliefs about the environment and the actor's beliefs. EPR is also limited by how discriminable the goals and plans under consideration are in the context of the observations and the model. An imperfect observer lacking complete observations may wrongly infer a plan or goal or fail to discriminate between a number of possible hypotheses. In particular, the quality of PR can suffer if observations are noisy (e.g., as a result of a faulty sensor) or missing altogether as observed and addressed by Sohrabi et al. [47]. Finally, the observer's beliefs about the actor's beliefs must be sufficiently complete and accurate for the observer to robustly infer the actor's plan and goal. We elaborate on this point in the rest of this section.

Recall that an EPR problem is cast from the perspective of the observer. Let $\mathcal{I}^*_{\text{Act}}$ and $\mathcal{D}^*_{\text{Act}}$ denote the actor's *actual* beliefs about the world and the set of action descriptions (as opposed to the observer's beliefs about the actor's beliefs), respectively. We say that a plan π achieves a goal G wrt the actor's actual beliefs if π is progressed in the context of $\mathcal{I}^*_{\text{Act}}$ and $\mathcal{D}^*_{\text{Act}}$.

Definition 3.5 (Adequacy). Given an EPR problem $R = \langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O\rangle$ and the actor's actual beliefs $Z = (\mathcal{I}^*_{\text{Act}}, \mathcal{D}^*_{\text{Act}})$, we say that $\mathcal{I}$ and $\mathcal{D}$ are **adequate** wrt (R, Z) *iff* for every goal $G \in \mathcal{G}$ and plan π that satisfies O, π achieves $B_{\text{Act}}G$ wrt the observer *iff* π achieves G wrt the actor's actual beliefs.

If the observer's beliefs about the actor's beliefs are adequate, then the observer can generate (wrt the actor) precisely all plans that satisfy O and achieve some $G \in \mathcal{G}$, which the actor can generate. In our example, if the observer's beliefs about the actor's beliefs are not adequate and it falsely believes B_{Act}*loc*(Party,Address1), then the actor's presumed plan (heading to Address2) might be incoherent to the observer wrt the goal *at*(Cake,Party).

Proposition 3.6. *Given an EPR problem $R = \langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O\rangle$ and the actor's actual beliefs $Z = (\mathcal{I}^*_{ACT}, \mathcal{D}^*_{ACT})$, let π be a plan (executable in $\mathcal{I}^*_{ACT}$) that satisfies O and achieves G wrt the actor's actual beliefs for some $G \in \mathcal{G}$. If $\mathcal{I}$ and $\mathcal{D}$ are adequate wrt (R, Z) then π is not incoherent wrt G.*

Proof sketch: suppose that the actor is following a plan π that is incoherent from the observer's perspective wrt goal G. It follows from the fact that $\mathcal{I}$ and $\mathcal{D}$ are adequate wrt (R, Z) and from the fact that π achieves G wrt the actor's actual beliefs, that π also achieves $B_{\text{Act}}G$ wrt the observer. This is in contradiction with the assumption that π is incoherent from the observer's perspective.

That is, since $\mathcal{I}$ and $\mathcal{D}$ can generate any plan π the actor can generate, any such plan will 'make sense' to the observer. It is therefore beneficial for the observer to strive to improve the adequacy of its beliefs about the actor's beliefs to better distinguish between ill-formed and incoherent plans. Indeed, an assistive observer will handle the two types of plans differently - assist the actor in achieving the goal if π is ill-formed (e.g., the observer helping Alice find the correct address) or strive to refine $\mathcal{I}$ and $\mathcal{D}$ (e.g., by querying the actor) if π is incoherent.

Inadequacy may be addressed in a myriad of ways, including the modulation of the observer's confidence in its predictions in the face of perceived irrationality (i.e., deviation from expected optimality) [29]. Masters and Sardina [29] identified that deviation from expected optimality may stem from a number of possible reasons including a deceptive actor and the observer's misconceptions regarding the actor's beliefs. Inadequacy of $\mathcal{I}$ and $\mathcal{D}$ falls under the latter case where, while the actor may indeed be irrational, the observed irrationality may be due to the inadequacy of the observer's beliefs about the actor's beliefs. While in our computation and experimentation we assume an omniscient observer (who thus has adequate beliefs about the actor's beliefs), future work can employ Masters and Sardina's *rationality measure* (that quantifies the agent's deviation from the observer's expectation of optimality) to allow the observer to self-modulate its confidence in its predictions.

In addition to self-modulating its confidence, the observer may strive to improve the adequacy of $\mathcal{I}$, $\mathcal{D}$, and the observations at its avail. To this end, the observer could, for instance, query the actor or sense the world in order to gain additional information, as is done in recent work on *active goal recognition* [45]. Further, Pollack discusses inference techniques by which an observer can infer an actor's beliefs (and plans) given dialogue utterances formed by the latter [39]. The observer may also refine its beliefs about the actor's beliefs by employing machine learning techniques (e.g., [26, 37]). Finally, When $\mathcal{D}$ is inadequate, the oberver may hold false beliefs about the actor's beliefs about the effects or preconditions of actions; Pereira et al. [36] explore goal recognition in such settings.

4 COMPUTATION

Recall that a solution to an EPR problem is a pair (π, G) where π is a plan that satisfies the sequence of observations O and $G \in \mathcal{G}$ is a goal. A number of different criteria have been proposed in the past to select a plan and goal that 'best' align with O, including the simplicity of the plan and the likelihood of a goal given the observations. Our general specification supports a diversity of criteria and computational realizations and here we propose one such realization of EPR as epistemic planning by appealing to the *plan recognition as planning* paradigm, proposed by Ramírez and Geffner [41]. Ramírez and Geffner transform the plan recognition problem into a planning problem, allowing for the use of off-the-shelf planning tools to solve the recognition problem. In what follows, we describe a transformation of EPR to epistemic planning; a way by which to compute a probability distribution over the set of possible goals given the transformed planning problem; and an algorithm that uses these techniques to compute a solution to an EPR problem.

4.1 Transformation to Epistemic Planning

Our transformation is inspired by the transformations proposed by Ramírez and Geffner [42] and by Sohrabi et al. [47] who later modified Ramírez and Geffner's approach. While Ramírez and Geffner cast the classical PR problem as a classical planning problem, we realize a computational solution to the EPR problem by appealing to epistemic planning and utilizing epistemic planners.

Intuitively, Ramírez and Geffner's approach compiles the observations in O into the planning domain, thereby forcing all generated plans that solve the transformed planning problem to *satisfy* O. In the EPR setting, we define a correspondence between a given EPR problem and an MEP problem by augmenting $\mathcal{D}$ with *explain* actions that allow the planner to explain the observations. Further, $\mathcal{P}$ is augmented with special predicates p_i, l_{α_i}, and p_{α_i} for each observation o_i that are made true when o_i is explained. These predicates ensure that the order of the observation sequence O is respected by all plans that solve the transformed MEP problem. $\mathcal{P}$ is also augmented with p_{init} which is set to be true initially. Formally, given an EPR problem $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O\rangle$ and some goal $G \in \mathcal{G}$, we create an MEP problem $\langle\langle \mathcal{P}', \mathcal{D}', Ag\rangle, \mathcal{I}', G'\rangle$ such that:

- $\mathcal{P}' = \mathcal{P} \cup \{p_i \mid (\alpha_i, \phi_i) \in O\} \cup \{l_{\alpha_i} \mid (\alpha_i, \phi_i) \in O\} \cup \{p_{init}\}$
- $\mathcal{D}' = \mathcal{D} \cup \mathcal{D}_{explain}$
  - $\mathcal{D}_{explain} = \{\langle l_{\alpha_i} \wedge p_{i-1}, \{(\top, p_{\alpha_i})\}\rangle \mid (\alpha_i, \phi_i) \in O\} \cup \{\langle \phi_i \wedge p_{\alpha_i}, \{(\top, p_i)\}\rangle \mid (\alpha_i, \phi_i) \in O\}$
- $\mathcal{I}' = \mathcal{I} \wedge p_{init}$
- $G' = B_{\text{ACT}}G \wedge p_m$, where o_m is the last observation in O

We add $(\top, l_{\alpha_i})$ and p_{i-1} to the conditional effects and precondition, respectively, of every action description in $\mathcal{D}$ that corresponds to an action that appears in an observation $(\alpha_i, \phi_i) \in O$ (where α_i is not empty). For every observation o_i, we augment $\mathcal{D}$ with two explain actions to ensure that α_i is accounted for before ϕ_i is explained. We do this since ϕ_i is assumed to be observed immediately *following* the execution of α_i. If α_i is empty, the preconditions of the corresponding explain action will be p_{i-1} rather than $l_{\alpha_i} \wedge p_{i-1}$. Similarly, if $\phi_i = \top$ then the preconditions of the corresponding explain action will be p_{α_i} rather than $\phi_i \wedge p_{\alpha_i}$. The precondition of the explain action in $\mathcal{D}_{explain}$ corresponding to the first observation in O is set to be p_{init}.

The order in which the observations are explained is enforced by the precondition p_{i-1} which only allows an observation $(\alpha_i, \phi_i) \in O$ to be explained after all the observations in O which precede it have been explained. Further, the transformation ensures that if G' is achieved then G is achieved wrt the actor and may (but need not) be achieved from the observer's perspective. Therefore, a plan that achieves G' may be ill-formed.

We use $\langle Q', \mathcal{I}', G'\rangle$ as shorthand for $\langle\langle \mathcal{P}', \mathcal{D}', Ag\rangle, \mathcal{I}', G'\rangle$. With the described correspondence, solutions to the transformed MEP problem, $\langle Q', \mathcal{I}', G'\rangle$, capture precisely the solutions to the corresponding MEP problem within the EPR problem that satisfy O, as stated by the following theorem.

THEOREM 4.1. *Given an EPR problem, $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O\rangle$, some goal $G \in \mathcal{G}$ and the corresponding transformed MEP problem, $\langle Q', \mathcal{I}', G'\rangle$, we have that:*
(1) If π is a sequence of actions that satisfies O and solves $\langle\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag\rangle, \mathcal{I}, B_{ACT}G\rangle$ then there exists a sequence of actions π' that solves $\langle Q', \mathcal{I}', G'\rangle$ such that π can be reconstructed straightforwardly from π' by removing the explain actions from π'.
(2) If π' is a sequence of actions that solves $\langle Q', \mathcal{I}', G'\rangle$ then there exists a sequence of actions π that solves $\langle\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag\rangle, \mathcal{I}, B_{ACT}G\rangle$ and satisfies O such that π can be reconstructed straightforwardly from π' by removing the explain actions from π'.

Proof sketch: Recall that the explain actions in $\mathcal{D}_{explain}$, together with the special predicates, ensure that the observations are explained in the correct order. Plans that solve the transformed MEP problem therefore satisfy the sequence of observations O. Further, the explain actions do not change the state of the world or the epistemic state of any agent. Thus, when removing the explain actions from a plan π that solves the transformed MEP problem, π solves $\langle Q', \mathcal{I}', B_{\text{ACT}}G\rangle$ (but no longer solves $\langle Q', \mathcal{I}', G'\rangle$) and also solves $\langle\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag\rangle, \mathcal{I}, B_{\text{ACT}}G\rangle$. This is because the only actions left in π after removing the explain actions are drawn from $\mathcal{D}$.

4.2 Computing $P(G|O)$

We build on the approach proposed by Ramírez and Geffner [42], using epistemic planners instead of classical planners, and compute the probability distribution over $\mathcal{G}$, $P(G|O)$, given the transformed MEP problem. To compute the probability distribution, two plans are generated for every goal $G \in \mathcal{G}$ - one that satisfies O (where $G' = B_{\text{ACT}}G \wedge p_m$) and one that does not (where $G' = B_{\text{ACT}}G \wedge \neg p_m$). We then define Δ as the cost difference between the costs of these two plans and use Δ to compute the probability of a goal. Formally, Bayes' Rule is used to compute $P(G|O) = \alpha P(O|G)P(G)$ where α is a normalization constant and $P(G)$ is the prior probability of G, which we assume in this work to be uniform across $\mathcal{G}$. Finally, assuming a Boltzmann distribution as in Ramírez and Geffner [42]:

$$P(O|G) \approx \frac{e^{-\beta\Delta}}{1 + e^{-\beta\Delta}} \qquad (1)$$

where β is a positive constant. Ramírez and Geffner assume a soft rationality postulate according to which G is a better predictor of O when Δ is smaller. Note that the actor's rationality is assumed wrt the *observer's* model of the former. When the observer is not omniscient, its beliefs may not be adequate, and the quality of inferences may consequently suffer (as is also the case when Δ is computed with sub-optimal planners).

4.3 Computing a Solution to an EPR Problem

Algorithm 1 describes how to obtain a solution to an EPR problem by leveraging off-the-shelf epistemic planners. In Line 3, for each goal $G \in \mathcal{G}$, the function TRANSFORMEPRTOMEP transforms an EPR problem R and a goal $G \in \mathcal{G}$ to an MEP problem, as described in Section 4.1. In Line 4, COMPUTEDELTA runs a planner twice on the transformed MEP problem - once with $G' = B_{\text{ACT}}G \wedge p_m$ and once with $G' = B_{\text{ACT}}G \wedge \neg p_m$. In Line 5, COMPUTEPROBABILITY uses the cost difference between the plans, Δ, to compute a posterior probability $P(G|O)$ for $G \in \mathcal{G}$, using Equation 1. In Line 7, we select the goal G^m with the highest posterior probability $P(G|O)$ (we randomly break ties between goals). In Line 8, the function RETRIEVEASSOCIATEDPLAN retrieves the plan π (forced to satisfy O) that was generated in Line 4 in order to compute Δ for G^m.

In both our computation and experimentation the observer is assumed omniscient and disjunctive beliefs are precluded. Further, for some action $a \in \mathcal{A}$, all corresponding action descriptions in $\mathcal{D}$ have the same precondition. Finally, the actor is assumed to only pursue one goal $G^* \in \mathcal{G}$ at a given time and we assume that a plan π that achieves G^* (wrt the actor) and satisfies O exists.

Algorithm 1

1: **procedure** SolveEPRProblem($\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O \rangle$) - Given an EPR problem R=$\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O \rangle$, return ($\pi$,$G$), where $G \in \mathcal{G}$ is a goal and π is a plan that satisfies O and achieves G wrt the actor.
2: **for each** $G \in \mathcal{G}$ **do**
3: $\langle Q', \mathcal{I}', G' \rangle \leftarrow$ TransformEPRToMEP(R, G)
4: $\Delta \leftarrow$ ComputeDelta($\langle Q', \mathcal{I}', G' \rangle$)
5: $P(G|O) \leftarrow$ ComputeProbability(Δ)
6: **end for**
7: $G^m \leftarrow \arg\max_{G \in \mathcal{G}}(P(G|O))$
8: $\pi \leftarrow$ RetrieveAssociatedPlan(G^m)
9: **return** (π, G^m)
10: **end procedure**

Algorithm 1 computes solutions from a subset of EPR solutions where the plan component π has the added property of achieving G wrt the actor, as stated by the following theorem.

Theorem 4.2. *Given an EPR problem, $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O \rangle$, and a sound and complete epistemic planning algorithm, Algorithm 1 returns a pair (π, G) where $G \in \mathcal{G}$ is a goal and π is a plan that satisfies O and achieves the goal G wrt the actor, if such a solution exists.*

Proof sketch: In Line 9, Algorithm 1 returns (π, G^m). π solves the transformed MEP problem corresponding to the given EPR problem and G^m, and it follows from Theorem 4.1 that π satisfies O. Further, From the construction of the corresponding MEP problem it follows that π achieves $B_{\text{ACT}}G^m$ wrt the observer and G^m wrt the actor.

Theorem 4.3. *The runtime complexity of Algorithm 1 given an EPR problem $\langle \mathcal{P}, \mathcal{A}, \mathcal{D}, Ag, \mathcal{I}, \mathcal{G}, O \rangle$ is $O(2 \cdot |\mathcal{G}| \cdot C) = O(|\mathcal{G}| \cdot C)$ where C is the runtime complexity of the chosen epistemic planner.*

The proof follows straightforwardly from the fact that the runtime complexity of Algorithm 1 is dominated by the two calls to an epistemic planner in Line 4 for each goal $G \in \mathcal{G}$. In contrast to Ramírez and Geffner's classical planning formalism, epistemic planning has been shown to be significantly more computationally expensive (e.g., [1]). Indeed, deciding whether a plan exists for an MEP problem in its general form has been shown to be an undecidable problem [5]. There are, however, a number of decidable and expressive fragments of epistemic planning (e.g., [9]). For example, when all preconditions are propositional and actions change only beliefs, the plan existence problem is in EXPSPACE [6]. In Section 6, we discuss the complexity of the various epistemic planners used in our experimentation to compute a solution for the transformed MEP problem, as well as the implications of employing a particular fragment of epistemic planning in the EPR setting.

5 EMPIRICAL EVALUATION

In this section, we present the results of our experimental evaluation. In our evaluation, we set out to (1) demonstrate that existing epistemic planners can straightforwardly be used to solve EPR problems and recognize epistemic goals pursued by an actor (2) compare the performance of existing epistemic planners in terms of computation time and their ability to assign the highest probability to the true goal the actor is pursuing and (3) demonstrate that the quality of inferences suffers when the observer has inadequate beliefs about the beliefs of the actor. While Algorithm 1 produces a solution to an EPR problem comprising both a goal G and a plan π that satisfies O, we focus here on the goal recognition task - inferring the actor's true goal $G^* \in \mathcal{G}$. To this end, we compute $P(G|O)$ in Line 5 of Algorithm 1 to determine the likelihood of goals in $\mathcal{G}$ given the sequence of observations and evaluate whether G^* was assigned the highest probability. We constructed and encoded a diversity of domains and ran them using the latest version of 3 off-the-shelf planners - RP-MEP [33], MEPK [18], and Efp [24].

Epistemic Planning Benchmarks. In the *Grapevine* (GV(n, m)) domain, n guests are in a villa with m rooms. Initially, each guest has her own secret to share with other guests. In this modified version, only one agent can freely move between the rooms and share her secrets with the guests in the room she is currently in. The goals involve some guests obtaining some (or no) secrets. Possible epistemic goals include misconception (an agent holds a false belief about someone else's belief) and a universal spread of information. In the *Selective-Communication* (SC(n,m)) domain, there are n agents distributed amongst m rooms along a corridor, and one agent can freely move between neighboring rooms. After an agent shares information, all agents in the room and in neighboring rooms know the information. The agent must first find out the information before being able to share it with other agents. Possible epistemic goals include agents knowing some or all of the information.

To create EPR problem instances, we populated a set of possible *epistemic* goals, $\mathcal{G}$, for each planning instance in the problem set, where $\mathcal{G}$ comprised the hidden goal (the true goal pursued by the actor) and 6 other possible goals ($|\mathcal{G}| = 7$ for all problem instances). We generated the observation sequence for each problem instance by sampling hidden optimal plans for a hidden goal concatenated to the MEP domain within the EPR problem. When sampling an action from optimal plans, the corresponding observation comprises both the action and the effects of the corresponding action description. We included additional agents in Ag to enable the actor to reason about their beliefs. We ran Algorithm 1 for each problem instance. In Line 4, to compute the Δ with RP-MEP, which encodes the MEP problem as a classical planning problem, we call the Fast Downward planner [16] with an admissible heuristic twice for each goal, configuring the planner to only compute optimal plans. MEPK and Efp are both optimal planners and we ran each planner twice for each goal in each problem instance to find Δ. We ran all planners on a 3.4GHz Intel Core i5 machine with 16 GB RAM since we were not able to run Efp on our more powerful machine.

Table 1 summarizes the results for the GV and SC domains, offering a comparison between the three different epistemic planners we used. Each row in the table is an average over 10 EPR problem instances, with a varying percentage of observations sampled from the hidden optimal plan: 10%, 40%, 70%, and 100%. We also vary d, the required depth of nested belief modalities in the problem instance. The T column represents the average time it took to run Algorithm 1 and solve the EPR problem instances. For example, it took RP-MEP, on average, 3.87 seconds to solve the $GV(2, 3)$ setting, with 70% of sampled observations and $d = 1$. The Q column represents the quality of the solution, i.e., whether or not the hidden goal was found to be most likely, and is computed using the probability $P(G|O)$ computed in Line 5 of Algorithm 1. For example, $Q = 0.89$ signifies that in 89% of problem instances, the hidden goal was amongst the goals found to be most likely. The variances ranged 0-1.37 and 2.12-3.91 for the Q and T values (for a particular row and planner), respectively. Timeout was 30 minutes. We could not run Efp on $GV(4, 5)$ problems due to modelling difficulties.

Satisficing Epistemic Planners. To compare with the optimal epistemic planners, we ran all GV and SC experiments with RP-MEP, coupled with a *satisficing* configuration of Fast Downward [16]. We also experimented with PG-Efp, a satisficing heuristic search planner that employs a heuristic derived from an epistemic planning graph [24]. In the case of RP-MEP, the satisficing planner was, on average, much faster than the optimal planner we used. However, the accuracy (Q) decreased significantly (particularly with incomplete observations) since the satisficing planner often generated highly suboptimal plans. In turn, the quality of inferences suffered as Δ was computed with low-quality plans, which differed greatly from the optimal ground truth. The results are reported in Table 1 (RP-MEP (S) column). As for PG-Efp, we observed a reduction in computation time in some cases (as reported by Le et al. [24]). Interestingly, the accuracy did not suffer greatly as the planner typically found optimal plans. Ramírez and Geffner [42] have shown that satisficing classical planners can be successfully used in PR to greatly reduce computation time, without significantly hurting accuracy. Thus, the EPR as epistemic planning approach will benefit from future research on satisficing epistemic planners.

With respect to our **first and second objectives**, the results present a comparison between existing epistemic planners and demonstrate that these planners can indeed be used to solve EPR problems and successfully recognize an actor's epistemic goals. Note that the performance, in seconds, of the three planners wrt d is consistent with Le et al.'s results. For instance, when $d = 3$ Efp's performance is not affected whereas RP-MEP slows down considerably. We will discuss some limitations of our approach and existing epistemic planners in the next section.

In Need of Assistance. The following scenario is inspired by Talamadupula et al.'s work, set in an urban search & rescue setting [49]. In our scenario, Eve (Act) is in need of assistance from either Alice or Bob (Obs). The three agents are each initially located in a different cell on a grid. Since Alice moved to a different location unbeknownst to Eve, the latter holds a false belief pertaining to Alice's location. Bob knows that Alice moved and also believes that Eve is not aware of the move. Next, Eve is observed to be heading from her initial location to some location which is on the way to where she believes Alice is. We modelled this scenario as an EPR problem and performed a set of simulations by varying the parameters of this problem (e.g., grid size and agent locations), resulting in 20 problems. $\mathcal{G}$ contains all cells on the grid. The observations were sampled from Eve's hidden plan which consists of her making her way to where she falsely believes Alice is (and thus Alice's incorrect location is Eve's hidden goal). We run this scenario once with Bob the observer holding adequate beliefs about Eve's beliefs and once with Bob holding inadequate beliefs (by modifying $\mathcal{I}$ such

Table 1: Comparison between three optimal epistemic planners in the Grapevine (GV) and Selective-Communication (SC) domains. Results in the RP-MEP (S) column generated using a satisficing planner. Each row describes averages over ten EPR problems, where the columns stand for % of actions sampled (%O), required depth of nested belief (d), avg time in seconds to solve problem instance (T), avg quality measuring fraction of problems where hidden goal is among the most likely (Q). L is the avg optimal plan length for the hidden goal. TO and n/a signify a timeout and inability to model the problem, respectively. $|\mathcal{G}| = 7$ in all problems.

			RP-MEP		MEPK		Efp		RP-MEP (S)	
	%O	d	T	Q	T	Q	T	Q	T	Q
$GV(2,3)\|L = 4$	10	1	1.72	1	0.12	1	1.49	1	0.84	0.37
	40	1	2.92	1	0.13	1	11.91	1	1.04	0.42
	70	1	3.87	1	0.14	1	37.82	1	1.57	0.63
	100	1	4.76	1	0.22	1	58.12	1	1.82	0.85
	10	3	549.24	1	0.31	1	1.49	1	60.42	0.34
	40	3	572.87	1	0.39	1	11.91	1	102.43	0.48
	70	3	591.43	1	0.41	1	37.82	1	96.49	0.69
	100	3	601.76	1	0.46	1	58.12	1	170.42	0.88
$GV(4,5)\|L = 11$	10	1	576.23	0.89	25.76	0.89	n/a	n/a	168.42	0.33
	40	1	584.86	1	54.43	1	n/a	n/a	152.89	0.39
	70	1	590.21	1	85.35	1	n/a	n/a	172.44	0.54
	100	1	598.19	1	TO	TO	n/a	n/a	162.91	0.71
	10	3	739.89	0.89	26.89	0.89	n/a	n/a	250.37	0.32
	40	3	753.91	1	57.12	1	n/a	n/a	248.11	0.45
	70	3	772.83	1	88.18	1	n/a	n/a	302.78	0.58
	100	3	801.79	1	TO	TO	n/a	n/a	296.54	0.74
$SC(8,4)\|L = 7$	10	1	67.12	0.85	0.36	0.85	324.89	0.85	5.42	0.33
	40	1	75.98	1	0.48	1	395.21	1	2.21	0.41
	70	1	77.63	1	0.61	1	457.21	1	1.94	0.64
	100	1	79.31	1	0.71	1	503.36	1	3.05	0.78
	10	3	912.54	0.85	9.54	0.85	324.89	0.85	256.24	0.35
	40	3	925.08	1	14.21	1	395.21	1	225.41	0.39
	70	3	949.21	1	18.48	1	457.21	1	249.28	0.65
	100	3	964.76	1	24.01	1	503.36	1	297.53	0.86

that Bob believes that Eve knows Alice's true location, when in fact she does not). For each problem instance and each Bob, we ran Algorithm 1 with RP-MEP to compute $P(G|O)$ over $\mathcal{G}$. As shown in Table 2, 'adequate' Bob was able to assign the highest probability to Eve's true goal (getting to where she believes Alice is) in most cases, compared to 'inadequate' Bob, who performed poorly.

With respect to our **third objective**, the results demonstrate that when Bob's beliefs about Eve's beliefs are inadequate, the quality of inferences indeed suffers. Further, suppose Bob is trying to infer whether Eve is looking for him or for Alice. 'Inadequate' Bob would infer that given the observations it is more likely that Alice is heading towards his location (one of the goals in $\mathcal{G}$) than towards Alice's *actual* location which is in the opposite direction of Eve's trajectory and is thus assigned a lower probability. However, 'adequate' Bob's beliefs about Eve's beliefs allow him to reason that the observations put Eve on the optimal path to where *she believes* Alice is located. Thus, Bob can reason that Eve is looking for Alice and that her presumed plan is ill-formed wrt this goal. Finally, we note that, more generally, an 'inadequate' observer can do much better or worse than 'inadequate' Bob, depending on the relevance of what they (do not) know.

Table 2: Comparison of an observer with adequate beliefs about Eve's beliefs and one with inadequate beliefs, on 20 EPR 'In Need of Assistance' problems. The values represent the percentage of problems in which Eve's true goal was assigned the highest probability given the observations, relative to the percentage of total observations (%O).

Observer Type \ %O	10	30
Observer w/ Inadequate Beliefs	10	15
Observer w/ Adequate Beliefs	**90**	**100**

6 DISCUSSION AND SUMMARY

In this work, we have introduced the notion of EPR, which appeals to a rich epistemic logic framework to model the observer in the PR setting, represent agent beliefs, and allow for the recognition of epistemic goals. We proposed a computational realization of EPR as epistemic planning that enables the use of existing planning tools. Finally, we performed an experimental evaluation of our approach on a set of EPR problems by utilizing existing epistemic planners.

There is a diverse body of research related to the ideas we have presented. PR research has mostly utilized (possibly hierarchical) plan libraries to best match a sequence of observations (e.g., [2, 15, 19]). Recent work, however, has aligned itself with the bottom-up approach to PR which views it as the reverse process of planning and leveraged advances in AI planning research by casting the PR problem as a planning problem (e.g., [42]). However, none of these works explicitly model the observer's mental state nor do they address epistemic goals. Work on BDI has also explored the PR problem (e.g., [10, 39, 46]) with a variety of applications, including air-combat modelling [43]. The limitation of these approaches, however, is that they require a specification of agent plans in advance. Further, these approaches did not appeal to planning techniques to generate plans. The work most related to ours is that of Talamadupula et al. which combines belief modelling with PR [49]. However, their work does not appeal to an epistemic logic, nor does it address the recognition of epistemic goals.

While the results of our experimentation are promising, they do not come without limitations. As discussed, the runtime complexity of Algorithm 1 is dominated by the two calls to the chosen epistemic planner. Huang et al.'s planner uses a satisfiability solving algorithm which has an exponential time complexity. In Le et al.'s planner, the size of the state grows exponentially. The encoding process in Muise et al.'s planner generates an exponential number of fluents when classically encoding the problem. Our approach calls an epistemic planner twice for each goal and will therefore benefit from advances in epistemic planning including faster computation and satisficing planners. Lastly, we wish to explore other epistemic planning paradigms that were not included in our experiments (e.g., Engesser et al. [11]; Hu et al. [17]; Fabiano [12]). Relatedly, a setting similar to epistemic planning is that of collaborative multi-agent planning under uncertainty with partial observability, which is typically modelled with Dec-POMDPs and [3] Qualitative Dec-POMDPs [7]. Synergies between the two paradigms will be explored.

As mentioned in the discussion of adequacy in Section 3, PR systems are limited by the veracity and completeness of the observer's beliefs. Even if the observer's beliefs about the actor's beliefs are adequate, the planner used may not be sound and complete, in which case it may not generate all plans the actor can generate. Indeed, a challenge of epistemic, conformant and contingent planning, is the issue of belief space representation to make planning computationally feasible, often at the expense of planner completeness and applicability. For example, conformant and contingent planners often appeal to belief state approximations such as 0-approximation [48] and related work (e.g., [34]). In epistemic planning, syntactic restrictions or approximations (e.g., PEKBs) may restrict applicability or limit completeness of planners and are traded off against scalability. For instance, if the depth of nested belief is restricted to 2, formulae such as $B_{John}B_{Alice}\phi$ may appear in the KB but $B_{John}B_{Alice}B_{Bob}\phi$ may not, which may prevent the observer from inferring the actor's plan or goal. Another such syntactic restriction is the exclusion of disjunctive belief, as is done in our computation and experiments and in PEKBs more generally. Additionally, Miller et al. [30] have shown that representing that an agent knows whether ϕ (a restricted form of disjunction) can facilitate inference about that agent's plans. It is therefore important, in the EPR setting, to consider the trade-off between computational feasibility and the expressivity of the chosen fragment of logic. Finally, future work will relax the assumption of observer omniscience and address the interesting computational challenges that arise in such settings.

Online Recognition and Runtime Optimization. In our experimentation, we focus on a setting where recognition is done post-hoc. Future work will explore computational approaches that are better geared towards an online recognition setting, where the observer is attempting to recognize a plan that is in-progress (e.g., [14, 50]). Further, we may utilize existing landmark-based approaches to goal recognition (e.g., [35, 50]) to compute a probability distribution over $\mathcal{G}$ given the classical planning problem that results from Muise et al.'s compilation. These approaches have been shown to be much faster than Ramírez and Geffner's approach, while achieving similar recognition accuracy [35].

Decentralized Multi-agent Setting. The EPR model in this work captures a fixed observer which is adequate in many settings such as tutoring systems and conversational systems. Like centralized vs distributed control, this model affords computational advantages over a more general setting. Our account of EPR could be extended to a decentralized multi-agent setting where each agent is both an actor and an observer and holds EPR capabilities, with a view to active collaboration or adversarial interactions.

Obfuscating Actor. In cases where the actor is aware of being observed, modelling their beliefs about the observer can be useful as these could affect the actor's behavior. For example, an actor keen on obfuscating its goal or plan might purposefully generate an ambiguous plan that is predicated on the actor's beliefs about the observer's beliefs [20, 22, 28]. Our specification can accommodate various approaches to plan and goal obfuscation (or legibility [8]).

ACKNOWLEDGEMENTS

We gratefully acknowledge funding from the Natural Sciences and Engineering Research Council of Canada (NSERC). We thank the anonymous reviewers for their comments and suggestions and Tara Gralnick for her helpful insights into agents Claire and Tara.

REFERENCES

[1] Guillaume Aucher and Thomas Bolander. 2013. Undecidability in Epistemic Planning. In *IJCAI-International Joint Conference in Artificial Intelligence-2013*.

[2] Dorit Avrahami-Zilberbrand, G Kaminka, and Hila Zarosim. 2005. Fast and Complete Symbolic Plan recognition: Allowing for Duration, Interleaved Execution, and Lossy Observations. In *Proc. of the AAAI Workshop on Modeling Others from Observations, MOO*.

[3] Daniel S Bernstein, Robert Givan, Neil Immerman, and Shlomo Zilberstein. 2002. The Complexity of Decentralized Control of Markov Decision Processes. *Mathematics of operations research* 27, 4 (2002), 819–840.

[4] Thomas Bolander. 2017. A Gentle Introduction to Epistemic Planning: The DEL Approach. *arXiv preprint arXiv:1703.02192* (2017).

[5] Thomas Bolander and Mikkel Birkegaard Andersen. 2011. Epistemic planning for single-and multi-agent systems. *Journal of Applied Non-Classical Logics* 21, 1 (2011), 9–34.

[6] Thomas Bolander, Martin Holm Jensen, and Francois Schwarzentruber. 2015. Complexity results in epistemic planning. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*. AAAI Press, 2791–2797.

[7] Ronen Brafman, Guy Shani, and Shlomo Zilberstein. 2013. Qualitative planning under partial observability in multi-agent domains. In *Twenty-Seventh AAAI Conference on Artificial Intelligence*.

[8] Tathagata Chakraborti, Anagha Kulkarni, Sarath Sreedharan, David E Smith, and Subbarao Kambhampati. 2019. Explicability? legibility? predictability? transparency? privacy? security? the emerging landscape of interpretable agent behavior. In *Proceedings of the International Conference on Automated Planning and Scheduling*, Vol. 29. 86–96.

[9] Tristan Charrier, Bastien Maubert, and François Schwarzentruber. 2016. On the Impact of Modal Depth in Epistemic Planning. In *IJCAI*. 1030–1036.

[10] Philip R Cohen, C Raymond Perrault, and James F Allen. 1981. Beyond question answering. *Strategies for natural language processing* (1981), 245–274.

[11] Thorsten Engesser, Thomas Bolander, Robert Mattmüller, and Bernhard Nebel. 2017. Cooperative epistemic multi-agent planning for implicit coordination. *arXiv preprint arXiv:1703.02196* (2017).

[12] Francesco Fabiano. 2019. Design of a Solver for Multi-Agent Epistemic Planning. *arXiv preprint arXiv:1909.08259* (2019).

[13] Ronald Fagin, Joseph Y Halpern, Yoram Moses, and Moshe Vardi. 2004. *Reasoning about knowledge.* MIT press.

[14] Richard G Freedman and Shlomo Zilberstein. 2017. Integration of Planning with Recognition for Responsive Interaction Using Classical Planners.. In *AAAI*. 4581–4588.

[15] Robert P. Goldman, Christopher W. Geib, and Christopher A. Miller. 1999. A New Model of Plan Recognition. In *UAI*. 245–254.

[16] Malte Helmert. 2006. The fast downward planning system. *Journal of Artificial Intelligence Research* 26 (2006), 191–246.

[17] Guang Hu, Tim Miller, and Nir Lipovetzky. 2019. What you get is what you see: Decomposing Epistemic Planning using Functional STRIPS. *arXiv preprint arXiv:1903.11777* (2019).

[18] Xiao Huang, Biqing Fang, Hai Wan, and Yongmei Liu. 2017. A general multi-agent epistemic planner based on higher-order belief change. In *IJCAI 2017, Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Australia, Sydney*.

[19] Henry A Kautz and James F Allen. 1986. Generalized Plan Recognition. In *AAAI*. 32–37.

[20] Sarah Keren, Avigdor Gal, and Erez Karpas. 2016. Privacy Preserving Plans in Partially Observable Environments. In *IJCAI*. 3170–3176.

[21] Filippos Kominis and Hector Geffner. 2015. Beliefs In Multiagent Planning: From One Agent to Many. In *ICAPS*. 147–155.

[22] Anagha Kulkarni, Siddharth Srivastava, and Subbarao Kambhampati. 2019. A unified framework for planning in adversarial and cooperative environments. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33. 2479–2487.

[23] Gerhard Lakemeyer and Yves Lespérance. 2012. Efficient Reasoning in Multiagent Epistemic Logics. In *ECAI*. 498–503.

[24] Tiep Le, Francesco Fabiano, Tran Cao Son, and Enrico Pontelli. 2018. EFP and PG-EFP: Epistemic forward search planners in multi-agent domains. In *Twenty-Eighth International Conference on Automated Planning and Scheduling*.

[25] Hector J Levesque. 1998. A completeness result for reasoning with incomplete first-order knowledge bases. In *KR*. 14–23.

[26] Lin Liao, Donald J Patterson, Dieter Fox, and Henry Kautz. 2007. Learning and inferring transportation routines. *Artificial Intelligence* 171, 5-6 (2007), 311–331.

[27] Yongmei Liu, Gerhard Lakemeyer, and Hector J Levesque. 2004. A Logic of Limited Belief for Reasoning with Disjunctive Information. In *KR*. 587–597.

[28] Peta Masters and Sebastian Sardina. 2017. Deceptive Path-Planning.. In *IJCAI*. 4368–4375.

[29] Peta Masters and Sebastian Sardina. 2019. Goal recognition for rational and irrational agents. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 440–448.

[30] Tim Miller, Paolo Felli, Christian Muise, Adrian Pearce, and Liz Sonenberg. 2016. 'Knowing whether' in proper epistemic knowledge bases. In *Thirtieth AAAI Conference on Artificial Intelligence*.

[31] Tim Miller and Christian J Muise. 2016. Belief Update for Proper Epistemic Knowledge Bases.. In *IJCAI*. 1209–1215.

[32] Christian Muise, Tim Miller, Paolo Felli, Adrian R Pearce, and Liz Sonenberg. 2015. Efficient reasoning with consistent proper epistemic knowledge bases. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 1461–1469.

[33] Christian J. Muise, Vaishak Belle, Paolo Felli, Sheila A. McIlraith, Tim Miller, Adrian R. Pearce, and Liz Sonenberg. 2015. Planning Over Multi-Agent Epistemic States: A Classical Planning Approach. In *AAAI*. 3327–3334.

[34] Hector Palacios and Héctor Geffner. 2007. From Conformant into Classical Planning: Efficient Translations that May Be Complete Too.. In *ICAPS*. 264–271.

[35] Ramon Fraga Pereira, Nir Oren, and Felipe Meneguzzi. 2017. Landmark-based heuristics for goal recognition. In *Thirty-First AAAI Conference on Artificial Intelligence*.

[36] Ramon Fraga Pereira, André Grahl Pereira, and Felipe Meneguzzi. 2019. Landmark-Enhanced Heuristics for Goal Recognition in Incomplete Domain Models. In *Proceedings of the International Conference on Automated Planning and Scheduling*, Vol. 29. 329–337.

[37] Ramon Fraga Pereira, Mor Vered, Felipe Meneguzzi, and Miquel Ramırez. 2019. Online Probabilistic Goal Recognition over Nominal Models. In *International Joint Conference on Artificial Intelligence 2019*. International Joint Conferences on Artificial Intelligence, 5547–5553.

[38] Ronald PA Petrick and Fahiem Bacchus. 2002. A Knowledge-Based Approach to Planning with Incomplete Information and Sensing.. In *AIPS*, Vol. 2. 212–222.

[39] Martha E Pollack. 1986. A model of plan inference that distinguishes between the beliefs of actors and observers. In *of the 24th annual meeting on Association for Computational Linguistics*. Association for Computational Linguistics, 207–214.

[40] David Premack and Guy Woodruff. 1978. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences* 1, 4 (1978), 515–526.

[41] Miquel Ramírez and Hector Geffner. 2009. Plan Recognition as Planning. In *IJCAI*.

[42] Miquel Ramírez and Hector Geffner. 2010. Probabilistic plan recognition using off-the-shelf classical planners. In *Proceedings of the Conference of the Association for the Advancement of Artificial Intelligence (AAAI 2010)*. 1121–1126.

[43] Anand S Rao and Graeme Murray. 1994. Multi-agent mental-state recognition and its application to air-combat modelling. In *of the Workshop on Distributed Artificial Intelligence (DAI-94)*.

[44] Charles F. Schmidt, NS Sridharan, and John L. Goodson. 1978. The plan recognition problem: An intersection of psychology and artificial intelligence. *Artificial Intelligence* 11, 1-2 (1978), 45–83.

[45] Maayan Shvo and Sheila A. McIlraith. 2020. Active Goal Recognition. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI-20)*.

[46] Michal P Sindlar, Mehdi M Dastani, Frank Dignum, and John-Jules Ch Meyer. 2008. Mental state abduction of BDI-based agents. In *Workshop on Declarative Agent Languages and Technologies*. Springer, 161–178.

[47] Shirin Sohrabi, Anton V Riabov, and Octavian Udrea. 2016. Plan Recognition as Planning Revisited.. In *IJCAI*. 3258–3264.

[48] Tran Cao Son and Chitta Baral. 2001. Formalizing sensing actions—a transition function based approach. *Artificial Intelligence* 125, 1-2 (2001), 19–91.

[49] Kartik Talamadupula, Gordon Briggs, Tathagata Chakraborti, Matthias Scheutz, and Subbarao Kambhampati. 2014. Coordination in human-robot teams using mental modeling and plan recognition. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2957–2962.

[50] Mor Vered, Ramon Fraga Pereira, Maurício Cecílio Magnaguagno, Gal A Kaminka, and Felipe Meneguzzi. 2018. Towards Online Goal Recognition Combining Goal Mirroring and Landmarks.. In *AAMAS*. 2112–2114.