An AI Planning-Based Approach to the Multi-Agent Plan Recognition Problem

Maayan Shvo *
University of Toronto
Toronto, Canada

Shirin Sohrabi
IBM Research
Yorktown Heights, USA

Sheila A. McIlraith
University of Toronto
Toronto, Canada

Abstract

Plan Recognition is the problem of inferring the goals and plans of an agent given a set of observations. In Multi-Agent Plan Recognition (MAPR) the task is extended to inferring the goals and plans of multiple agents. Previous MAPR approaches have largely focused on recognizing team structures and behaviors, given perfect and complete observations of the actions of individual agents. However, in many real-world applications of MAPR, observations are unreliable, i.e., unexplainable or missing; they are often over properties of the world rather than actions, and the observations that are made may not be explainable by the agents' goals and plans. Moreover, the actions of the agents could be temporal or concurrent. In this paper, we address the problem of MAPR with temporal actions and with observations that can be unreliable. To this end, we propose a multi-step compilation technique that enables the use of AI planning for the computation of the probability distributions of plans and goals, given observations. In addition, we propose a set of novel benchmarks that enable a standard evaluation of solutions that address the MAPR problem with temporal actions and such observations. We present results of an experimental evaluation on this set of benchmarks, using several temporal and diverse planners.

1 Introduction

Plan recognition (PR) – the ability to recognize the plans and goals of agents from observations – is useful in a myriad of applications including intelligent user interfaces, conversational agents, intrusion detection, video surveillance, and now increasingly in support of human-robot and robot-robot interactions (e.g., (Carberry 2001)). Originally conceived in the context of single agent plan recognition (e.g., (Cohen *et al.* 1981), (Schmidt *et al.* 1978), (Kautz and Allen 1986), (Charniak and Goldman 1993)), recent work has turned to the more complex task of Multi-Agent Plan Recognition (MAPR). In MAPR, the goals and/or plans of multiple agents are hypothesized, based upon observations of the agents, providing a richer paradigm for addressing many of the applications noted above. Early work in this area (e.g., (Banerjee *et al.* 2010)) limited observations to activity-sequences, and focused the recognition task on the identification of dynamic team structures and team behaviors, relative to a predefined plan library.

While this formulation is effective for certain classes of problems, it does not capture important nuances that are evident in many real-world MAPR tasks. To this end, we provide in this paper an enriched characterization of MAPR that provides support for a richer representation of the capabilities of agents and the nature of observations. In particular, we support (1) differing skills and capabilities of individual agents; (2) agent skills and actions that are durative or temporal in nature (e.g., washing dishes or other durative processes (cf. (Fox and Long 2003))); (3) observations with respect to the state of the system; such observations range over fluents rather than over actions as actions may not be directly observable but rather inferred via the changes they manifest; (4) observations that are missing or unexplainable (i.e. cannot be accounted for by agents' actions).

Our approach to addressing this problem is to conceive the computational core of MAPR as a planning task, following in the spirit of the single-agent characterization of *plan recognition as planning* proposed by Ramírez and Geffner 2009. This contrasts with much of the previous work on MAPR which requires explicit plan libraries; and while the work done by Zhuo et al. 2012 replaces explicit plan libraries with sets of action models, it does not make use of AI planning. In our work, the conception of *MAPR as planning* enables the leveraging of recent advances in multi-agent planning as exemplified by the planners that participated in the 2015 Competition of Distributed and Multiagent Planners (CoDMAP) (e.g., (Crosby *et al.* 2014; Muise *et al.* 2015b)), as well as advances in temporal planning (e.g., (Benton *et al.* 2012), and in the generation of diverse plans (e.g., (Nguyen *et al.* 2012)). Furthermore, the use of AI planning enables us to capture both the possible interaction between the different agents, and the temporal aspects of a domain. Importantly, the use of AI planning enables not only the recognition of goals given observations, but also the recognition of plans given observations.

To realize MAPR as planning, we propose a two-step compilation process that takes a MAPR problem as input. We first compile away the multi-agent aspect of the problem and then we compile away the observations. The resulting planning problem is temporal, has temporal actions and temporal constraints; hence, temporal or makespan-sensitive planners can be applied to generate plans that are then post-processed to yield a solution to the original MAPR problem.

*The work was performed during an internship at IBM

We propose three different approaches to generating high-quality MAPR results, evaluating them experimentally. Using these approaches, we are able to compute the probability distributions of plans and goals, given observations. Our approach also enables us to zoom in on a particular subset of agents and recognize their goals, thus offering a greater resolution when solving the MAPR problem.

The main contributions of this paper are: (i) a formalization of the MAPR problem with unreliable observations over fluents, and actions that are temporal or durative in nature; (ii) characterization of MAPR as planning via a two-step compilation technique that enables the use of temporal AI planning on the transformed planning problem; (iii) three approaches to computing the probability distributions of goals and plans given the observations; (iv) a set of novel benchmarks that will allow for a standard evaluation of solutions to the MAPR problem; (v) experimental evaluation and comparison of our proposed techniques on this set of benchmarks using several temporal and diverse planners.

2 Problem Definition

In this section, we review basic definitions, and introduce the multi-agent plan recognition problem with temporal actions and its solution. We begin with an example that will serve to illustrate different concepts throughout the paper.

Running Example: Let us consider an example, illustrated in Figure 1, taken from the International Planning Competition (IPC) Depots domain; in this domain, there are two different types of agents: hoist operators and truck drivers. Each agent may have their own goal, yet a common goal might be shared by the agents and distributed amongst them; in some cases, it is not possible to solve the planning problem of each agent separately since resources are shared between agents and their activities are interdependent and complementary. For example, a truck driver must wait for a hoist operator to load a crate onto the truck, before being able to drive it to its designated location. The goal of the agents is moving a set of crates between different locations.

Given a MAPR problem where (1) actions are temporal and may occur concurrently (e.g., two truck driver agents driving at the same time); (2) agents have different skill sets (e.g., truck driver agents drive and hoist operator agents lift); and (3) observations could be unreliable (e.g., a faulty sensor might give an incorrect location of a truck), the task is to recognize the goals and the plans of the agents given the observations (for example, the dashed and solid red lines represent the hypothesized plans of the red truck, given the observations). This requires taking into consideration all manner of interaction between agents, in addition to the temporal aspects of the domain, and cannot be achieved by breaking apart the MAPR problem into many single agent plan recognition problems. Instead, our approach transforms the MAPR problem into a temporal planning problem whose plans and makespans approximate the probability distributions of goals and plans given the observations.

We consider a planning problem, a tuple $P^c = (F, A, I, G)$, where F is a finite set of fluent symbols, A is a set of actions, $I \subseteq F$ defines the initial state, and $G \subseteq F$ defines the goal state. Each action $a \in A$ is associated with a precondition, pre(a), add effects, add(a), and delete effects, del(a). A state, s, is a set of fluents that are true. An action $a \in A$ is *executable* in a state s if pre(a) $\subseteq s$. The successor state is defined as $\delta(a, s) =((s \setminus$ del(a)) $\cup$ add(a)) for the executable actions. The sequence of actions $\pi = [a_1, ..., a_n]$ is executable in s if the state $s' = \delta(a_n, \delta(a_{n-1}, \ldots, \delta(a_1, s)))$ is defined. Moreover, π is the solution to the planning problem P^c if it is executable from I and $G \subseteq \delta(a_n, \delta(a_{n-1}, \ldots, \delta(a_1, I)))$. Next, we modify the above definition, to include temporal actions as defined in (Fox and Long 2003).

A planning problem with temporal actions is a tuple $P^t = (F, A, I, G)$, where F, I, and G are defined as before, and A is a set of temporal actions. Each $a \in A$ is associated with a duration, d(a), precondition at start, pre$_s(a)$, precondition over all, pre$_o(a)$, precondition at end, pre$_e(a)$, add effects at start, add$_s(a)$, add effects at end, add$_e(a)$, delete effects at start, del$_s(a)$, and delete effects at end, del$_e(a)$. The semantics of a temporal action is often given using two non-temporal actions "start" and "end". Here we provide a similar semantics that instead uses "start" and "end" states. A temporal action $a \in A$ is *executable* in a state s$_{start}$, ending in state s$_{end}$ if pre$_s(a) \subseteq$ s$_{start}$ and pre$_e(a) \subseteq$ s$_{end}$. The resulting states s$_{start'}$ and s$_{end'}$ are defined as s$_{start'}$ = ((s$_{start} \setminus$ del$_s(a)$) $\cup$ add$_s(a)$) and s$_{end'}$ = ((s$_{end} \setminus$ del$_e(a)$) $\cup$ add$_e(a)$). Note that s$_{end}$ comes after s$_{start'}$. Additionally, the overall precondition, pre$_o(a)$ must hold in every state between s$_{start'}$ and s$_{end}$. The solution to P^t is a set of action-time pairs, allowing actions to occur concurrently, where each action is executable, and the goal G holds in the final state. The makespan of the solution is the total time that elapses between the beginning of the first action and the end of the final action. Also as is mentioned, it is possible for two actions to occur concurrently. Next, we extend the multi-agent planning problem with temporal actions.

Definition 1 (MAPP with Temporal Actions) *A Multi-Agent Planning Problem (MAPP) with temporal actions is a tuple $P^m = (F, \{A_i\}_{i=1}^N, I, G)$, where F, I are defined as before, G is the goal of the multi-agent problem, achieved by N agents, each with their own set of temporal action descriptions, A_i, $1 \leq i \leq N$.*

Note that the notion of concurrency amongst actions is modeled via temporal actions. This stands in contrast to much past research (e.g., (Brafman and Domshlak 2008; Crosby *et al.* 2014)) which used joint actions and defined concurrency constraints over them. The use of joint actions to model concurrent actions performed by multiple agents is restricting in that a single agent cannot perform two actions concurrently. Use of temporal actions allows concurrency of a single agent's actions as well as actions of different agents. Note that in this work we assume that all agents know and operate over one (completely specified) state of the world, with identical knowledge about it; this allows for the straightforward encoding of the multi-agent planning problem as a single-agent temporal planning problem. However, in general, where agents have differing, incom-

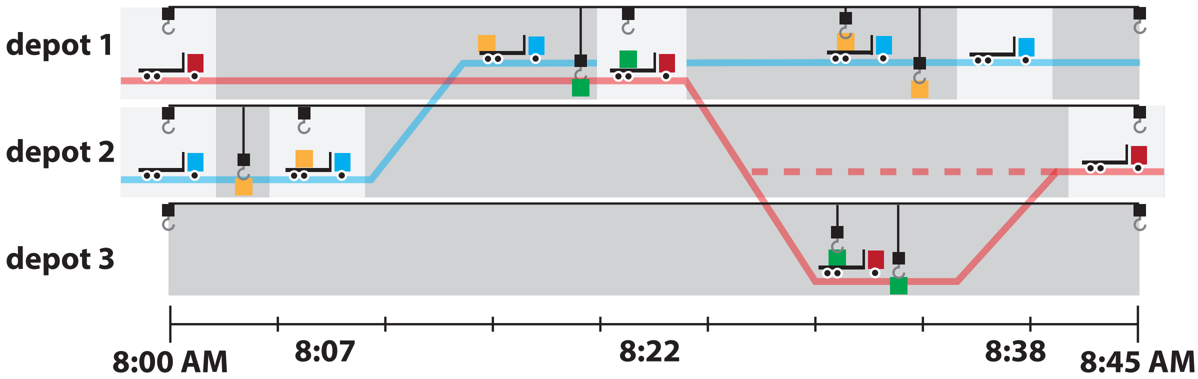

Figure 1: A timeline illustrating a 45-minute timeframe for the Depots domain, with 2 trucks drivers and 3 hoist operators. The y-axis is the three depot locations; the x-axis is the time line. The white areas indicate the observations. The lines represent alternative possible sequences of actions given the observations (color image).

plete knowledge of the world, and perhaps different levels of rationality, establishing the correspondence between single-agent and multi-agent planning is more challenging, and this is left for future work.

Next, we define the plan recognition problem with temporal actions, as well as unexplainable and missing observations, adapting the definitions of Sohrabi et al. 2016, where quality as measured by cost is used instead of action durations to approximate the probability values.

Definition 2 (PR Problem with Temporal Actions) *A plan recognition problem with temporal actions is a tuple $P^r = (F, A, I, O, \mathcal{G}, \text{PROB})$, where F, I, are defined as before, A is a set of temporal actions as defined earlier, $O = [o_1, ..., o_m]$ is the sequence of observations, where $o_k = (f_k, t_k)$, $1 \leq k \leq m$, $f_k \in F$ is the observed fluent, t_k is the time f_k was observed, and $\forall o_i, o_j$, if $i < j$ then $t_i < t_j$. $\mathcal{G}$ is the set of possible goals G, $G \subseteq F$, and PROB is the probability of a goal, $P(G)$, or the goal priors.*

Definition 3 (Unexplainable/Missing Observations) *Given an observation sequence O and a plan π for a particular goal G, an observation $o = (f, t)$ in O is said to be unexplainable (aka noisy), if f is a fluent that does not arise as the consequence of any of the actions a_i from the plan π for G. In contrast, an observation $o' = (f', t')$ is said to be missing from O, if o' is not in the sequence O and f' is added by at least one of the executed actions $a_i \in \pi$.*

In this paper, we consider sequences of observations where each observation $o_i \in O$ is an observable fluent, with a timestamp that indicates when that fluent was observed. We focus on observations that range over fluents rather than over actions, as actions may not be directly observable but rather inferred via the changes they manifest. If two observable fluents are observed at the same time, we increase the timestamp of one by an arbitrarily small duration. Note, we assume that the act of observing an observation is instantaneous and adheres to the order in which the observable fluent appears in O. To illustrate, in Figure 1, ((at *redTruck depot1*), 08:00) is a possible observation in O.

Also note that both missing and unexplainable observations belong to the class of unreliable observations. To address the unexplainable observations, Sohrabi *et al.* 2016 modifies the definition of satisfaction of an observation sequence by an action sequence introduced in (Ramírez and Geffner 2010) to allow for observations to be left unexplained. Given an execution trace and an action sequence, an observation sequence is said to be satisfied by an action sequence and its execution trace if there is a non-decreasing function that maps the observation indices into the state indices as either explained or discarded. Hence, observations are all considered, while some can be left unexplained. The solution to the plan recognition problem, P^r, is two probability distributions, the probability of plans given observations, $P(\pi|O)$, and the probability of goals given observations, $P(G|O)$.

Next, we put everything together and define the problem we address in this paper.

Definition 4 (MAPR Problem with Temporal Actions) *The Multi-Agent Plan Recognition (MAPR) problem with temporal actions is described as a tuple $P = (F, \{A_i\}_{i=1}^N, I, O, Z, \mathcal{G}, \text{PROB})$, where F is a finite set of fluents, A_i is a set of temporal actions for agent i, $1 \leq i \leq N$, $I \subseteq F$ defines the initial state, $O = [o_1, ..., o_m]$ is the sequence of observations, where $o_k = (f_k, t_k)$, $1 \leq k \leq m$, $f_k \in F$ is the observed fluent, t_k is the time f_k was observed, Z is a set of agents (each element in Z corresponds to an index between 1 and N), $1 \leq |Z| \leq N$, $\mathcal{G}$ is the set of possible goals, $G \in \mathcal{G}$, pertaining to the set of agents Z, $G \subseteq F$, PROB is the prior probability of a goal, $P(G)$.*

Given a MAPR problem with temporal actions, P, a solution to P is in the form of two probability distributions. The first is the probability of plans given the observations, $P(\pi|O)$, where each π is a plan that achieves a goal $G \in \mathcal{G}$, satisfies the observation sequence, O, and involves at least

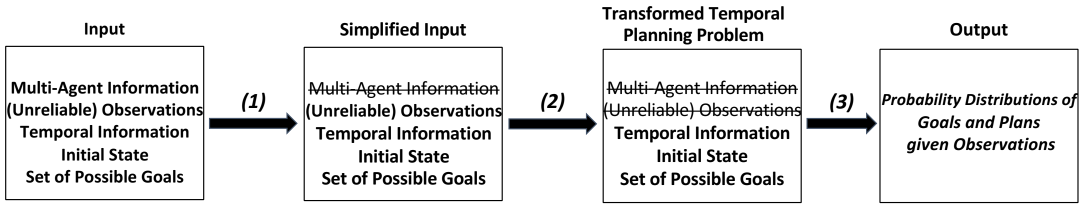

Figure 2: A pipeline showing our proposed compilation approach: transforming the original MAPR problem with temporal actions and unreliable observations into a plan recognition problem (1), a transformation step that compiles away the observations (2), allowing the use of temporal planning to compute a solution to the MAPR problem (3).

one action performed by an agent in Z. The second distribution is the probability of goals given the observations, $P(G|O)$, where each G assigned a non-zero probability is a goal achieved by a plan in the first distribution. To illustrate, consider a scenario where Ann, Bob, and Carol (A, B, C) are cooking dinner (agent names represent agent indices). C is working exclusively on dessert and has no interaction with A and B. A is making pasta sauce (which requires vegetables) and cooking the pasta and B is preparing the salad. B cuts all the vegetables and in so doing contributes to A's making of the sauce. In the case where $Z = \{A, B, C\}$, the goals might include made-dessert, made-salad, cooked-pasta, made-sauce. In the case where $Z = \{A, B\}$, we only keep goals that A or B could contribute to (i.e., remove made-dessert); where $Z = \{A\}$, we remove made-salad, and where $Z = \{B\}$, we remove cooked-pasta but not made-sauce from the set of goals.

In the case where Z includes all agents, the goals assigned a non-zero probability are those achieved by actions carried out by any agent; in the case where Z is restricted to a subset of agents, the attributed goals are only those achieved by plans which involve an agent in Z. Note, given a particular set of agents, Z, the solution characterizes the goals of the collective and does not distinguish goals of individuals. Our MAPR solution is generous in its attribution of a goal to individual agents, attributing to them not only goals solely achieved by them, but also goals they have contributed to; further, if Z contains a collection of agents, then goals are attributed to the collective in Z without further refinement to individual agents. In addition, each goal $G \in \mathcal{G}$ is a set of fluents that must hold in the final state. Our MAPR formalization allows us to zoom in on a particular set of agents, Z, and recognize the goals pursued by that set of agents. Thus, this formalization offers greater resolution given multiple instances of an MAPR problem, P, that only differ in Z, by way of enabling the recognition of goals and plans of different sets of agents. Note that while the agents may or may not be working cooperatively towards shared or independent goals, our formalization of the MAPR problem, in the current work, makes no assumptions about the cooperative nature of agents.

3 Transformation

In this section, we describe a two-step compilation technique that allows the use of temporal planning on the MAPR problem. That is, we first transform the given MAPR problem as defined in Definition 4 into a plan recognition problem with temporal actions; second, we transform the plan recognition problem into a temporal planning problem; finally, we use temporal planning to compute the solution to the MAPR problem, namely the probability distributions of plans and goals given observations, in keeping with the previous plan-recognition-as-planning approaches. The compilation pipeline is shown in Figure 2.

3.1 Transformation to Plan Recognition Problem

To transform the original MAPR problem with temporal actions to a single agent plan recognition problem with temporal actions, we compile away the multi-agent information by using special predicates that keep track of an agent's access to fluents and objects; every object o and agent i in the domain are assigned a corresponding fluent. For an agent i to be allowed to execute an action on object o, a precondition must be met, in which the corresponding fluent holds. To address the temporal aspect, the introduced action precondition is defined such that it meets the specifications of a temporal action; action durations are left unchanged. This approach is similar to that of Muise *et al.* 2015a, which maps multi-agent planning problems to single agent planning problems, thus enabling the use of single agent classical planners.

3.2 Transformation to Temporal Planning

Next, we compile away the observations, so that the plan recognition problem can be solved using the plan-recognition-as-planning approaches. These approaches view the plan recognition problem as an inverse planning problem, in that the goals and plans of the agents are not known to the system, and the goal of the transformed planning problem becomes explaining the given observations. There are several ways to compile away the observations, depending on the nature of the given observations. For example, if the observations are actions then one can take the approach described by Ramírez and Geffner 2009. Observations can also be compiled away following Haslum and Grastien 2011 using the "advance" action that ensures the observation order is preserved; another paper that addresses the compilation of observations is (Keren *et al.* 2016), where a goal recognition design problem is compiled into a classical planning problem and observations, which are over the agent's actions, are compiled into the transformed planning problem. In this paper, however, observations are defined over the fluents, so we will follow the technique proposed in (Sohrabi *et al.* 2016), which extends Ramírez and Geffner's approach by addressing unexplainable and missing observations. To incorporate a temporal aspect into the compilation process, our work replaces the notion of cost with that of duration, and compiles the observations into temporal actions that are part of the transformed temporal planning domain.

The transformation compiles away observations, using special predicates for each fluent in the observation sequence O, while ensuring that their order is preserved. The transformation ensures that observation o_1 with timestamp t_1 will be considered (explained or discarded) before observation o_2 with timestamp t_2, where $t_1 < t_2$; o_1, as explained previously, will appear before o_2 in the observation sequence O. To address the unexplainable observations, the set of actions, A, is augmented with a set of "discard" and "explain" actions for each observation o_i in the observation sequence, O, with a penalty for the discard action. We set the penalty by defining a high duration to the "discard" action, whereas in Sohrabi *et al.* 2016 the penalty was set by defining a high cost to the "discard" action. This penalty serves to encourage the planner to explain as many observations as possible. We also update the duration of the original actions, by adding a constant duration to each action; this is the penalty for the possible missing observations, which encourages the planner to use as few unobserved actions as possible. While these penalties artificially inflate the makespan of the plans, we are able to post-process these plans by removing the extra actions and updating the durations of the actions. To ensure that at least one of the given goals $G \in \mathcal{G}$ is achieved, and allow the use of a diverse planner that finds a set of plans, a special predicate "done" in addition to the corresponding predicate for the final fluent in the observation sequence are added to the goal of this transformed planning problem. In addition, we add an action for each goal $G \in \mathcal{G}$ with precondition g (the fluents corresponding to goal G), and add effect "done" to the set of actions. Hence, the goal of the transformed planning problem is set such that all observations are considered and achieve at least one of the goals $G \in \mathcal{G}$ is achieved.

Note, after solving the transformed single agent temporal planning problem, we are able to straightforwardly rewrite the solution, a single-agent temporal plan, as a multi-agent temporal plan, such that we can attribute the different actions in the plan to the corresponding agents. By so doing, we are able to hypothesize about the plan of any given set of agents, given the observations. For example, the blue truck's hypothesized plan, given the observations, is represented by the blue line in Figure 1. In order to apply the approach proposed in (Ramírez and Geffner 2010), as well as our proposed hybrid approach, we modify the transformation discussed above to not include the "done" predicate, as a new planning problem will be generated for each goal separately. In addition, the "discard" actions are removed for our proposed approach that is based on (Ramírez and Geffner 2010) as this approach does not address the unexplainable observations by discarding them.

Theorem 1 *Given a MAPR problem with temporal actions, $P = (F, \{A_i\}_{i=1}^N, I, O, Z, \mathcal{G}, \textsc{prob})$, as defined in Definition 4, where $|Z| = N$ (Z, here, is the set of all agents), and the corresponding transformed temporal planning problem $P' = (F', A', I', G')$ as described above, for all $G \in \mathcal{G}$, if π is a plan for the planning domain $(F, \{A_i\}_{i=1}^N, I)$ and goal G, then there exists a plan π' for the corresponding planning problem, P', such that the plan π can be constructed straightforwardly from π' by associating the actions in π' with the corresponding agents, removing the extra actions (i.e., discard, explain, and goal actions) and updating the duration of the remaining actions such that $d(\pi') = d(\pi) + M + (b_2 \cdot D)$, where $d(\pi)$ is the makespan of the plan π, M is the cumulative incurred penalty for missing observations, $b_1 \leq M \leq b_1 \cdot |\pi'|$ ($M = 0$ if $|\pi'| = 0$), $|\pi'|$ is the number of actions in π', D is the number of discard actions in π', and b_1 and b_2 are positive coefficients that express weights to the importance of missing and unexplainable observations, respectively.*

Proof is based on the fact that the extra actions (i.e., explain, discard, and goal) only preserve the ordering amongst the observations and do not change the state of the world. The duration of the actions in the transformed planning problem, P', incorporates the objective function that includes the original duration of the actions, as well as the penalty incurred for the missing and unexplainable observations. We add b_1 to the duration of all original actions, to account for missing observations; b_2 is the duration of the discard action, for the unexplainable observations. The duration of the explain and goal actions is 0. As will be explained in section 4, the makespans of plans in the transformed planning problem map to $V(\pi)$, which is used to approximate $P(O|\pi)P(\pi|G)$; thus, the probability distributions, $P(G|O)$ and $P(\pi|O)$, can be computed using these makespans.

4 Computation

In this section, we lay out our approaches to computing a solution to the MAPR problem, as described in Definition 4, namely the probability distributions of plans and goals, given observations. We begin by presenting our three approaches to computing the probability distribution of goals given the transformed planning problem, as described in the previous section, involving all N agents ($|Z| = N$). We then present an approach to computing the probability distribution of goals, involving a subset of the agents ($1 \leq |Z| < N$); this approach first pre-processes the MAPR problem, so that one of the three aforementioned approaches may be applied to the modified problem.

4.1 Computing $P(G|O), G \in \mathcal{G}, |Z| = N$

The first approach (Delta) is based on finding, for each of the different goals, the delta between the costs of two plans, one that explains the observations and one that does not; this method is a modification of the goal recognition approach, proposed in (Ramírez and Geffner 2010). The second approach (Diverse) computes the probability distribution of goals by finding a set of diverse plans, that serves as a representative approximation of the distribution of plans that satisfy the observations and achieve one of the possible goals ($P(\pi|O)$); it is a modification of the proposed approach in (Sohrabi *et al.* 2016). The third approach (Hybrid) is a combination of the two previous approaches, in that it computes a set of plans for each of the goals. Note that the Diverse and Hybrid approaches both compute the probability distribution of plans given observations in order to compute $P(G|O)$, while the Delta approach is not capable of doing so.

Approach 1 : Delta Given the transformed temporal planning problem, this approach computes the probability distribution of goals given observations, $P(G|O)$, by running the planner twice for each goal, once with the observations, and once without. More formally, $P(G|O)$ is computed using Bayes' Rule as:

$$P(G|O) = \alpha P(O|G)P(G) \tag{1}$$

where α is a normalization constant and $P(G)$ is PROB or the goal priors. The cost (or makespan) difference, or Δ, is defined as the difference in the makespan of the optimal plan that achieves G and O, and the makespan of the optimal plan that achieves G but not O. Assuming a Boltzmann distribution, $P(O|G)$ is defined as:

$$P(O|G) \approx \frac{e^{-\beta\Delta}}{1 + e^{-\beta\Delta}} \tag{2}$$

where β is a positive constant. This approach assumes that the agent pursing goal G is more likely to follow cheaper plans and that the probability that the agent is pursing a plan for goal G is dominated by the probability that the agent is pursing one of the most likely plans for goal G; hence, it only computes one plan for each setting of the problem.

Approach 2 : Diverse Given the transformed temporal planning problem, this approach computes an approximation to the probability distribution of plans as well as goals, given the observation, by running a diverse temporal planner on the transformed temporal planning problem. In particular, it first computes $P(\pi|O)$ as follows:

$$\begin{aligned} P(\pi|O) &= \beta P(O|\pi)P(\pi) \\ &= \beta P(O|\pi) \sum_G P(\pi|G)P(G) \\ &= \beta P(O|\pi)P(\pi|G)P(G) \end{aligned} \tag{3}$$

where β is a normalizing constant that depends on $P(O)$ only, and $P(G)$ is $\text{PROB}(G)$. Note, we assume that only one goal is being pursued and $P(\pi|G)$ is 0 for the action sequences π that are not plans for G. $P(O|\pi)P(\pi|G)$ is approximated as follows:

$$P(O|\pi) \cdot P(\pi|G) \approx 1 - \frac{\beta' V(\pi)}{\sum\limits_{\pi'' \in \Pi} V(\pi'')} \tag{4}$$

where β' is a positive constant, Π is a sampled set of plans that satisfy the observations and achieve at least one of the goals $G \in \mathcal{G}$; $V(\pi)$, which respects the objective function as mentioned in section 3, is the makespan of the plan that is the solution to the transformed planning problem and is equal to the sum of the original duration of the actions plus M plus b_2 times D. Coefficients b_1 (incorporated in M) and b_2 are used to give weights to the importance of the original actions together with the potential of having missing observations and unexplainable observations, respectively. D is the number of "discard" actions in π and $b_1 \leq M \leq b_1 \cdot |(\pi)|$ (see Theorem 1).

Using Bayes rule, the probability distribution of goals given observations is then computed by a summation over all values of $P(\pi|O)$ for the sampled set of plans, Π, that achieve G and satisfy O, and a subsequent normalization of the summation values (more details in (Sohrabi *et al.* 2016)).

$$P(G|O) = \sum_{\pi \in \Pi} P(\pi|O) \tag{5}$$

The set of plans Π is computed using diverse planning, where the objective is to find a set of plans m that are at least d distance away from each other. The solution to the diverse planning problem, (m, d), is a set of plans Π, such that $|\Pi| = m$ and $min_{\pi,\pi' \in \Pi}\ \delta(\pi, \pi') \geq d$, where $\delta(\pi, \pi')$ measures the distance between plans. Several techniques exist for computing the set of diverse plans (e.g., (Bryce 2014; Roberts *et al.* 2014; Srivastava *et al.* 2007; Coman and Muñoz-Avila 2011)); in this paper, we use LPG-d (Nguyen *et al.* 2012), the diverse extension of a local search-based planner LPG (Gerevini *et al.* 2003). Note that there is a large space of plans that achieve G and satisfy O and computing all of them is not practical; diverse planning is used as a means to approximate the probability distribution over these plans, by sampling a set of plans from this space.

The set of sampled plans is found by instructing the diverse planner to find m plans for the transformed temporal planning problem that satisfy the observation sequence and achieve at least one of the goals $G \in \mathcal{G}$.

Approach 3: Hybrid In order to take advantage of both previous approaches, we propose a hybrid approach in which we use a temporal planner to compute a smaller set of plans for each of the different goals. After merging the sets of plans, we are able to compute the probability distribution of goals, just as we did in the second approach. However, the Hybrid approach forces the planner to compute a set of plans for each of the goals, rather than allowing it to choose the goal that is shortest to reach. Thus, each of the possible goals is assigned at least one representative plan when computing the probability distribution over the different goals.

4.2 Computing $P(G|O), G \in \mathcal{G}, 1 \leq |Z| < N$

In the previous section, we considered three approaches to computing the probability distribution of goals given observations, involving all agents. In Algorithm 1, we propose an algorithm that allows us to use the same approaches when Z is restricted a subset of the agents. In step 1, $P(\pi|O)$ is computed by instructing the planner to find a set of plans that satisfy O and achieve at least one of the goals $G \in \mathcal{G}$ (see Section 4.1). In step 2, we remove actions from π that are not related to agents in Z; we use the same notion of interaction as defined in (Brafman and Domshlak 2008) (i.e., two agents are interacting if one agent's actions affect the

Algorithm 1: Computing the probability distribution of goals given observations, involving a set of agents Z

Input: MAPR problem $P = (F, \{A_i\}_{i=1}^{N}, I, O, Z, \mathcal{G}, \text{PROB})$
Output: Probability distribution over $G \in \mathcal{G}$, $P(G|O)$

1: For P, find a plan π with the highest posterior probability, $P(\pi|O)$, by applying the compilation process to P and using Equations 3 and 4.
2: Let π_Z be a partial plan based on π such that each action $a \in \pi_Z$ is either associated with at least one of the agents in Z or is associated with an agent that is interacting with an agent in Z as defined in (Brafman and Domshlak 2008).
3: Let O_Z be the sequence of observations based on π_Z (the effects of each action $a \in \pi_z$ serve as observations in O_Z).
4: Let $P' = (F, A', I, O_Z, Z, \mathcal{G}, \text{PROB})$ such that $A' = \{A_i\}_{i=1}^{Z} \cup \{a \mid a \in \pi_Z\}$.
5: Apply the compilation process to P' and follow one of the approaches described in Section 4.1 to compute $P(G|O)$, approximated as $P(G|O_Z)$, for each $G \in \mathcal{G}$.

functionality of the other agent). In the final step of the algorithm, we apply one of our three proposed approaches to P' and compute the probability distribution of goals involving **all** agents, relative to P'; the correctness of the algorithm follows from Theorem 1, where a correspondence is shown between an MAPR problem, where $|Z| = N$, and the transformed planning problem, allowing us to compute the probability distributions of goals and plans.

5 Experimental Evaluation

In this section, we present the results of our experimental evaluation; first, we evaluate the goal recognition capabilities of our three proposed computational approaches, involving all agents, where $|Z| = N$; second, we evaluate the recognition of individual goals, where $|Z| = 1$, using Algorithm 1; finally, we discuss the plan recognition capabilities of our MAPR approach.

To evaluate our MAPR approach, we used a temporal planner, LPG-TD (Gerevini *et al.* 2004), for the delta approach, the hybrid approach and to compute the joint plan in step 1 of Algorithm 1, and a diverse planner, LPG-d (Nguyen *et al.* 2012), for the diverse approach. We chose these planners as we were able to run them successfully, using the transformed planning problem as input. The other planners we have tested, (e.g., POPF2 (Coles *et al.* 2010), OPTIC (Benton *et al.* 2012)), either timed out on most problem instances, or did not accept the transformed planning problem as input. Note, the results for the diverse approach were obtained by running LPG-d once for each problem. LPG-TD was run once for each goal, i.e. $|\mathcal{G}|$ times, for the hybrid approach, $2 \times |\mathcal{G}|$ times for the delta approach and once in step 1 of Algorithm 1, to find $P(\pi|O)$. We used a timeout of 30 minutes and ran our experiments on dual 16-core 2.70 GHz Intel(R) Xeon(R) E5-2680 processor with 256 GB RAM. For the LPG-d planner we used a $(10, 0.2)$ setting of (m, d), since this setting performed best; 10 plans that are at least 0.2 distance away from each other. For the coefficients, we set b_1 to be the maximum of all action durations in the domain, and b_2 to be ten times b_1; thus, discarded observations are penalized more heavily than missing observations. The ratio between b_1 and b_2 was set based on experimentation.

Appr	Depots		Zeno		Rovers		Satellites	
	R	U	R	U	R	U	R	U
Delta	15/20	81/220	9/18	83/198	2/19	52/209	18/19	198/209
Diverse	15/20	146/220	8/18	90/198	11/19	154/209	14/19	173/209
Hybrid	15/20	160/220	14/18	152/198	19/19	207/209	19/19	204/209

Table 1: Comparison of the number of problems solved by our three proposed approaches, where $|Z| = N$.

In this paper, we address a combination of elements that has not been addressed by previous research; hence, we create, for evaluation purposes, a set of novel benchmarks, based on the International Planning Competition (IPC) domains and the Competition of Distributed and Multiagent Planners (CoDMAP), namely Rovers (a collection of rovers navigate a planet surface, finding samples and communicating them back to a lander), Depots (trucks transport crates between depots and then the crates must be stacked onto pallets at their destinations by hoist operators), Satellites (a collection of observation tasks carried out by multiple satellites, each equipped differently), and ZenoTravel (transportation of people between cities in planes, using different modes of movement). The original domains each have separate temporal and multi-agent versions and are not plan recognition problems. We modify the domains to create benchmark problems for the MAPR problem with temporal actions.

To construct the MAPR problems, we compute a plan that is a solution to the original planning problem. From this plan, we sample actions in order to construct O, the sequence of observations, while keeping track of the goal used in the original planning problem (i.e., ground truth goal). The effects of these actions then serve as the fluents in O. Additional goals were created manually for each problem instance to populate the set of possible goals; the goals were created with approximately equal prior likelihoods. Overall, the set of possible goals for each problem in the benchmarks consisted of 4 goals, involving all agents ($|Z| = N$). For each generated problem in the benchmarks, we created up to 3 sub-problems where $|Z| = 1$, in which we recognize the goals of individual agents, each consisting of 4 manually generated possible goals and where each sub-problem focuses on a different agent. The ground truth goal of each agent was determined based on the solution to the original planning problem, given the actions of the individual agent in the plan. In order to introduce missing observations, we created several variations of each problem that did not include the full observation sequence, by randomly selecting 10%, 40%, 70% and 100% of the observations in O.

Figure 3 shows the summary of the results when evaluating our three proposed MAPR approaches on goal recognition. Approach 1 is the delta approach, approach 2 is the diverse approach and approach 3 is the hybrid approach. Each domain consists of 16-20 problems; the problems vary in difficulty, thus the more difficult problems are computationally more complex. In addition, we have experimented by adding a number of extra observations, i.e., introduced

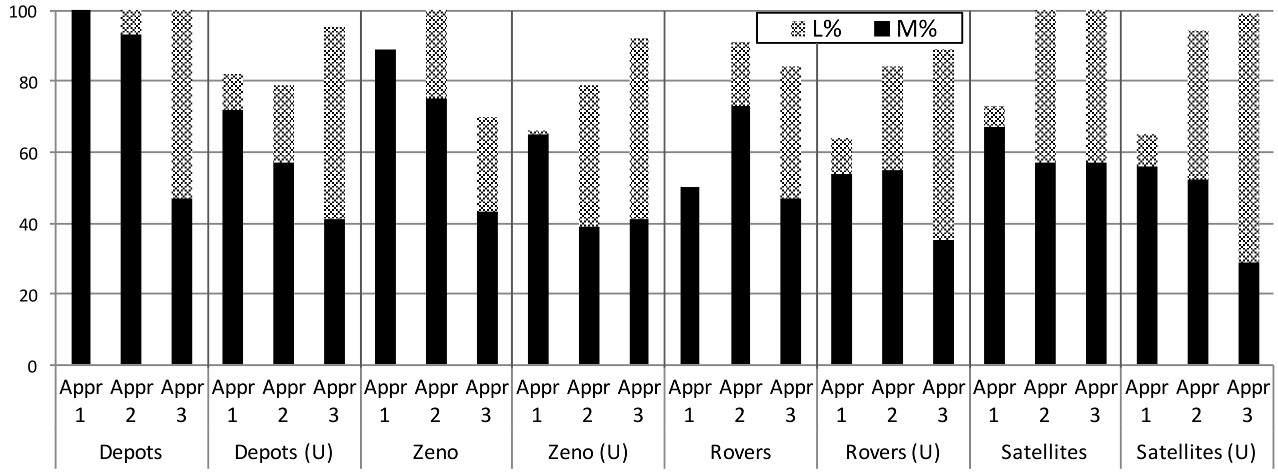

Figure 3: Comparison of our proposed approaches for recognizing a goal: (1) Delta, (2) Diverse, and (3) Hybrid, where $|Z| = N$.

noise; there are two levels of noise, one of which adds 12% extra, possibly unexplainable observations relative to the number of original observations, while the other adds the same percentage of noise, only this time relative to the size of the ground truth plan. The figure presents the results for each of the four domains, with and without the introduction of unreliable observations. U signifies that the results are an average over all cases where unreliable observations were introduced in a specific domain. We average over instances which were successfully solved before the timeout.

To evaluate the coverage and accuracy of the different approaches, we compute the average percentage of instances in which the ground truth goal was deemed Most and Less likely, i.e., whether or not the ground truth goal was assigned the highest posterior probability given the observations. These values, M and L, are shown respectively in the lower (solid black) and upper portions of the bars in the figure. The overall value of M and L, sum of the most and less likely percentages, indicates the goal recognition coverage for that method, and is expressed by the total height of each bar. The most likely goals are chosen relative to that particular approach (i.e., goals with the highest posterior probability) and the less likely goals are those goals with greater than 0.03 posterior probability.

The results in Figure 3 pertain to the $|Z| = N$ case (i.e. involving all agents) and show that approach 1, on average, does best (i.e., highest M value) across all domains when observations are reliable and no noise is introduced. The results also show that on average, approach 3 achieves the best coverage, i.e., the total height of the bars, across all domains, and also manages to successfully solve more problems than the other approaches. The total number of problems solved by each approach is shown in Table 1; U, as before, signifies cases where unreliable observations were introduced, whereas R signifies cases where they were not.

Note that the sheer amount of observations caused some problem instances to become computationally challenging, and often led to the system timing out; this can explain why the results are, in some cases, worse when less unreliable observations are introduced (i.e. less missing observations and hence larger observation sequences). Additionally, since the extra observations are added randomly, in some cases the observations are unexplainable, while in other cases, it is possible for the system to explain the extra observations by computing very long plans. This, combined with the the ground truth plans being sub-optimal, can account for some of the unexpected results, for example in cases where introduced noise does not hurt performance. The nature of the domains, e.g., interchangeable objects or inconsequential order of action execution also account for the different results.

We have also experimented with MAPR problems where $|Z| = 1$ (i.e. recognizing the goals of individual agents), using Algorithm 1 and the Delta approach. With no unreliable observations, our approach assigned the highest probability to the ground truth goal in the **Depots** domain in 93% of problem instances; **ZenoTravel** - 92%; **Rovers** - 100% and **Satellites** - 97%. Further experimental evaluation has shown that we cannot apply our three proposed approaches directly to MAPR problems where $1 \leq |Z| < N$; Algorithm 1, as explained in section 4.2, crucially pre-processes the MAPR problem before applying one of the three proposed computational approaches.

Finally, our approach, by conceiving the computational core of MAPR as a planning task, also allows us to recognize the plans of the agents. Since the problem addressed in the paper incorporates a temporal element, we are able to hypothesize about each agent's actions at any point along some timeline, given the observations. This plan recognition capability, utilized in Algorithm 1 in order to compute π in step 1, contributed to the promising results of the individual goal recognition experimentation, as described above. We plan to conduct further experimentation, testing both the robustness and scalability of our approach.

6 Related Work and Discussion

There exists a body of work on multi-agent systems. However, they often do not address temporal actions (e.g., (Kominis and Geffner 2015; Muise *et al.* 2015b; Bisson *et al.* 2015)). The closest to our work is the work of (Crosby *et al.* 2014), however, they do not model durations and the concurrency constraints are over objects. They also do not address the plan recognition problem. As for the plan recognition problem, it has been addressed, in many forms, by previous work (e.g., (Banerjee *et al.* 2015; Zhuo *et al.* 2012; Kominis and Geffner 2015; Sukthankar *et al.* 2014)). However, the focus and the problem addressed are different. In particular, they either do not address temporal actions, or do not make use of AI planning. In (Argenta and Doyle 2017), an approach to the MAPR problem is proposed which also makes use of AI planning; however, they focus on observations which are over actions rather than fluents and do not address temporal actions or unexplainable observations.

In this paper, we address the problem of MAPR with temporal actions and unreliable observations. To this end, we provide a formal characterization of the MAPR problem with temporal actions, and then propose a multi-step compilation technique that enables the use of AI planning for the computation of the probability distributions of goals and plans, given observations. In addition, we propose a set of novel benchmarks that allow for a standard evaluation of solutions that address the MAPR problem. We present results of an experimental evaluation of our approach on this set of benchmarks, using several temporal and diverse planners; the experimentation addresses both the recognition of goals involving all agents, and the recognition of goals of individual agents. To the best of our knowledge, we are the first to propose this problem and provide a solution for it. The merit of this paper is that it provides a way to solve an important class of MAPR problems that could not previously be addressed. It does so by leveraging and augmenting a combination of ideas from single agent plan recognition and multi-agent planning. Solving this variation of the MAPR problem, with unreliable observations and temporal actions, is paramount to the applicability of a MAPR approach to many real-world instantiations of the MAPR problem. Furthermore, our approach allows for much needed expressivity in the MAPR domain, while also providing the foundation for incorporation of various important and interesting aspects of MAPR, including, for example, agents with varying and limited knowledge of the state of the world and with differing physical and even cognitive capabilities. Finally, by enabling us to zoom in on a particular subset of agents and recognize their goals, our proposed formalization offers greater resolution when solving the MAPR problem.

Our work addresses different elements, namely temporal actions, a multi-agent setting and unreliable observations; this offers greater depth to our inferences regarding the goals and plans of agents. However, this also makes the problem addressed in this paper quite complex, rendering its computational solution expensive. Recent work (E-Martín *et al.* 2015) suggests an approach that propagates cost and interaction information in a plan graph, which are then used to estimate probabilities of goals. This proposed method might be enhanced and exploited here to reduce computation time. Further, we have attempted to apply the techniques in (Jiménez *et al.* 2015) so as to compile away the temporal aspect of the problem, thus transforming it to a single agent classical planning problem. However, this approach did not scale well, causing many of our experiments to time out. Finally, our work enables the application of a MAPR approach to previously unaddressed problems, by modeling them in planning domains. By enabling the use of existing temporal planners, one can choose the planner that works best for their domain and compute a solution to their MAPR problem.

Acknowledgements

We gratefully acknowledge Karin Niemantsverdriet's assistance with the design and creation of figures; and the Laboratory for Analytic Sciences (LAS) and the Natural Sciences end Engineering Research Council of Canada (NSERC) for funding this work. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of NSERC, the LAS and/or any agency or entity of the United States Government.

References

Chris Argenta and Jon Doyle. Probabilistic multi-agent plan recognition as planning (p-maprap): Recognizing teams, goals, and plans from action sequences. In *International Conference on Agents and Artificial Intelligence*, pages 575–582, 2017.

Bikramjit Banerjee, Jeremy Lyle, and Landon Kraemer. Multi-agent plan recognition: Formalization and algorithms. In *Proc. of the 24th National Conference on Artificial Intelligence (AAAI)*, 2010.

Bikramjit Banerjee, Jeremy Lyle, and Landon Kraemer. The complexity of multi-agent plan recognition. *Autonomous Agents and Multi-Agent Systems*, 29(1):40–72, 2015.

J. Benton, Amanda Coles, and Andrew Coles. Temporal planning with preferences and time-dependent continuous costs. In *Proc. of the 22nd Int'l Conference on Automated Planning and Scheduling (ICAPS)*, 2012.

Francis Bisson, Hugo Larochelle, and Froduald Kabanza. Using a recursive neural network to learn an agent's decision model for plan recognition. In *Proc. of the 24th Int'l Joint Conference on Artificial Intelligence (IJCAI)*, pages 918–924, 2015.

Ronen I. Brafman and Carmel Domshlak. From one to many: Planning for loosely coupled multi-agent systems. In *Proc. of the 18th Int'l Conference on Automated Planning and Scheduling (ICAPS)*, pages 28–35, 2008.

Daniel Bryce. Landmark-based plan distance measures for diverse planning. In *Proc. of the 24th Int'l Conference on Automated Planning and Scheduling (ICAPS)*, pages 56–64, 2014.

Sandra Carberry. Techniques for plan recognition. *User Modeling User-Adapted Interaction*, 11(1-2):31–48, 2001.

Eugene Charniak and Robert P Goldman. A bayesian model of plan recognition. *Artificial Intelligence*, 64(1):53–79, 1993.

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

Amanda Coles, Andrew Coles, Maria Fox, and Derek Long. Forward-chaining partial-order planning. In *Proc. of the 20th Int'l Conference on Automated Planning and Scheduling (ICAPS)*, pages 42–49, 2010.

Alexandra Coman and Hector Muñoz-Avila. Generating diverse plans using quantitative and qualitative plan distance metrics. In *Proc. of the 25th National Conference on Artificial Intelligence (AAAI)*, pages 946–951, 2011.

Matthew Crosby, Anders Jonsson, and Michael Rovatsos. A single-agent approach to multiagent planning. In *Proc. of the 21st European Conference on Artificial Intelligence (ECAI)*, pages 237–242, 2014.

Yolanda E-Martín, María D. R-Moreno, and David E. Smith. A fast goal recognition technique based on interaction estimates. In *Proc. of the 24th Int'l Joint Conference on Artificial Intelligence (IJCAI)*, pages 761–768, 2015.

Maria Fox and Derek Long. PDDL2.1: An extension to PDDL for expressing temporal planning domains. *Journal of Artificial Intelligence Research*, 20:61–124, 2003.

Alfonso Gerevini, Alessandro Saetti, and Ivan Serina. Planning through stochastic local search and temporal action graphs in LPG. *Journal of Artificial Intelligence Research*, 20:239–290, 2003.

Alfonso Gerevini, Alessandro Saetti, and Ivan Serina. LPG-TD: a fully automated planner for PDDL 2.2 domains. In *14th Int'l Conference on Automated Planning and Scheduling, system demo session (ICAPS-04)*, 2004.

Patrik Haslum and Alban Grastien. Diagnosis as planning: Two case studies. In *Int'l Scheduling and Planning Applications woRKshop (SPARK)*, pages 27–44, 2011.

Sergio Jiménez, Anders Jonsson, and Héctor Palacios. Temporal planning with required concurrency using classical planning. In *Proceedings of the 25th International Conference on Automated Planning and Scheduling (ICAPS)*, 2015.

Henry A Kautz and James F Allen. Generalized plan recognition. In *Proc. of the 5th National Conference on Artificial Intelligence (AAAI)*, pages 32–37, 1986.

Sarah Keren, Avigdor Gal, and Erez Karpas. Privacy preserving plans in partially observable environments. In *Proc. of the 25th Int'l Joint Conference on Artificial Intelligence (IJCAI)*, 2016.

Filippos Kominis and Hector Geffner. Beliefs in multiagent planning: From one agent to many. In *Proc. of the 25th Int'l Conference on Automated Planning and Scheduling (ICAPS)*, pages 147–155, 2015.

Christian Muise, Nir Lipovetzky, and Miquel Ramirez. MAP-LAPKT: Omnipotent multi-agent planning via compilation to classical planning. In *Competition of Distributed and Multiagent Planners*, 2015.

Christian J. Muise, Vaishak Belle, Paolo Felli, Sheila A. McIlraith, Tim Miller, Adrian R. Pearce, and Liz Sonenberg. Planning over multi-agent epistemic states: A classical planning approach. In *Proc. of the 29th National Conference on Artificial Intelligence (AAAI)*, pages 3327–3334, 2015.

Tuan Anh Nguyen, Minh Binh Do, Alfonso Gerevini, Ivan Serina, Biplav Srivastava, and Subbarao Kambhampati. Generating diverse plans to handle unknown and partially known user preferences. *Artificial Intelligence*, 190:1–31, 2012.

Miquel Ramírez and Hector Geffner. Plan recognition as planning. In *Proc. of the 21st Int'l Joint Conference on Artificial Intelligence (IJCAI)*, pages 1778–1783, 2009.

Miquel Ramírez and Hector Geffner. Probabilistic plan recognition using off-the-shelf classical planners. In *Proc. of the 24th National Conference on Artificial Intelligence (AAAI)*, 2010.

Mark Roberts, Adele E. Howe, and Indrajit Ray. Evaluating diversity in classical planning. In *Proc. of the 24th Int'l Conference on Automated Planning and Scheduling (ICAPS)*, pages 253–261, 2014.

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

Shirin Sohrabi, Anton Riabov, and Octavian Udrea. Plan recognition as planning revisited. In *Proc. of the 25th Int'l Joint Conference on Artificial Intelligence (IJCAI)*, 2016.

Biplav Srivastava, Tuan Anh Nguyen, Alfonso Gerevini, Subbarao Kambhampati, Minh Binh Do, and Ivan Serina. Domain independent approaches for finding diverse plans. In *Proc. of the 20th Int'l Joint Conference on Artificial Intelligence (IJCAI)*, pages 2016–2022, 2007.

Gita Sukthankar, Christopher Geib, Hung Hai Bui, David V. Pynadath, and Robert P. Goldman. *Plan, Activity, and Intent Recognition*. Morgan Kaufmann, Boston, 2014.

Hankz Hankui Zhuo, Qiang Yang, and Subbarao Kambhampati. Action-model based multi-agent plan recognition. In *Proc. of the 26th Annual Conference on Neural Information Processing Systems (NIPS)*, pages 377–385, 2012.