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

Maayan Shvo[1★(✉)], Shirin Sohrabi[2], and Sheila A. McIlraith[1]

[1] Department of Computer Science, University of Toronto, Canada
{maayanshvo, sheila}@cs.toronto.edu
[2] IBM Research, Yorktown Heights, NY, USA
ssohrab@us.ibm.com

Abstract. Multi-Agent Plan Recognition (MAPR) is the problem of inferring the goals and plans of multiple agents given a set of observations. While previous MAPR approaches have largely focused on recognizing team structures and behaviors, given perfect and complete observations, in this paper, we address potentially unreliable observations and temporal actions. 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. We present results of an experimental evaluation on a novel 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-machine and machine-machine interactions. Originally conceived in the context of single agent plan recognition (e.g., [1]), 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., [2]) 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

★ The work was performed during an internship at IBM durative processes (cf. [3])); (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 [4]. This contrasts with much of the previous work on MAPR which requires explicit plan libraries. 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. 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. 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.

2 Problem Definition

In this section, we review basic definitions, and introduce the multi-agent plan recognition problem with temporal actions and its solution.

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 is a finite set of fluent symbols, $I \subseteq F$ defines the initial state, and G is the goal of the multi-agent problem, achieved by N agents, each with their own set of temporal action descriptions, A_i.*

Each temporal action $a \in A_i$, as defined in [3], is associated with a duration, $\mathrm{d}(a)$, precondition at start, $\mathrm{pre}_s(a)$, precondition over all, $\mathrm{pre}_o(a)$, precondition at end, $\mathrm{pre}_e(a)$, add effects at start, $\mathrm{add}_s(a)$, add effects at end, $\mathrm{add}_e(a)$, delete effects at start, $\mathrm{del}_s(a)$, and delete effects at end, $\mathrm{del}_e(a)$. The semantics of a temporal action is often given using two non-temporal actions "start" and "end". Additionally, the overall precondition, $\mathrm{pre}_o(a)$ must hold in every state in between. The solution to P^m 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.

Next, we define the plan recognition problem with temporal actions, as well as unexplainable and missing observations, adapting the definitions of Sohrabi et al. [5], where quality as measured by cost is used instead of action durations.

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. To address the unexplainable observations, Sohrabi et al. [5] modifies the definition of satisfaction of an observation sequence by an action sequence introduced in [4] 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. Next, we 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 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.

3 Transformation

In this section, we briefly describe our multi-step compilation technique compilation technique that allows the use of temporal planning on the MAPR problem. The pipeline consists of transforming a MAPR problem as defined in Definition 4 into a single agent plan recognition problem with temporal actions, and a transformation step that compiles away the observations, allowing the use of temporal planning to compute the posterior probabilities of goals.

To transform the original MAPR problem with temporal actions to a single agent PR 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 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. We also add extra actions, "explain" and "discard" for each observation with a penalty to the "discard" action 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. 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 $t_1 < t_2$. Finally, in order to allow the use of diverse planning, the goal of the transformed planning problem is set such that all observations are considered and at least one of the goals $G \in \mathcal{G}$ is achieved.

Theorem 1. *Given a MAPR problem with temporal actions, $P = (F, \{A_i\}_{i=1}^{N}, I, O, Z, \mathcal{G}, \text{PROB})$, as defined in Def. 4, where $|Z| = N$, 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 problem $(F, \{A_i\}_{i=1}^{N}, I, G)$, then there exists a plan π' for the corresponding planning problem, P', such that the plan π can be constructed straightforwardly from π'.*

Proof is based on the fact that the extra actions only preserve the ordering amongst the observations and do not change the state of the world. 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 briefly describe 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. For further details, we refer the reader to [6].

Delta - This approach is based on finding, for each of the 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 [4]. The delta is found by running the planner twice for each goal.

Diverse - This approach 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 [5]. The set of diverse plans, which serves as an approximation to the probability distribution of plans and goals, given O, is found by running a diverse temporal planner on the transformed planning problem.

Hybrid - This approach is a combination of the two previous approaches, in that it computes a set of plans for each goal. In order to take advantage of both previous approaches, we propose an approach in which we use a temporal planner to compute a smaller set of plans for each of the goals. After merging the sets of plans, we are able to compute the probability distribution of goals, just as we did in the Diverse approach.

5 Experimental Evaluation

In this section, we briefly present the results of our experimental evaluation. For further details, we refer the reader to [6]. To evaluate our MAPR approach, we used a temporal planner, LPG-TD [7], for the delta approach and the hybrid approach, and a diverse planner, LPG-d [8], for the diverse approach. We have created, for evaluation purposes, a set of novel benchmarks, based on International Planning Competition (IPC) domains, namely Rovers, Depots, Satellites, and ZenoTravel. We modified the domains to create benchmark problems for the MAPR problem with temporal actions. In addition, we have introduced missing observations, by creating several variations of each problem that did not include the full observation sequence. Lastly, we have introduced noise by adding extra random observations, with two different levels of noise.

To evaluate the coverage and accuracy of our approaches on the goal recognition task, 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. Our results (shown in [6]) show that the Delta approach, on average, does best (i.e., deems the ground truth goal most likely) across all domains when observations are reliable and no noise is introduced. Further, on average, over all our problems, when unreliable observations were not introduced, **Delta** deemed the ground truth goal most likely in 77% of cases, **Diverse** - 75% of cases, and **Hybrid** - 49% of cases.

The total number of problems solved by each approach is as follows: **Delta** - 458/912; **Diverse** - 611/912; **Hybrid** - 790/912. Note that approach 1 manages to solve the least amount of problems, on average, compared to the other approaches. We plan to conduct further experimentation, testing both the robustness and scalability of our approach.

6 Discussion

The merit of this paper is that it provides a way to solve an important class of Multi-Agent Plan Recognition 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 class of problems, with unreliable observations and temporal actions, is paramount to the applicability of a MAPR approach to many real-world instantiations of the problem. Furthermore, our formalization allows for much needed expressivity, while also providing the foundation for incorporation of various important and interesting aspects of the problem, including, for example, agents with varying and limited knowledge of the state of the world and with differing physical and even cognitive capabilities. 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.

References

1. Schmidt, C.F., Sridharan, N., Goodson, J.L.: The plan recognition problem: An intersection of psychology and artificial intelligence. AIJ **11**(1-2) (1978) 45–83
2. Banerjee, B., Lyle, J., Kraemer, L.: Multi-agent plan recognition: Formalization and algorithms. In: AAAI. (2010)
3. Fox, M., Long, D.: PDDL2.1: An extension to PDDL for expressing temporal planning domains. JAIR **20** (2003) 61–124
4. Ramírez, M., Geffner, H.: Probabilistic plan recognition using off-the-shelf classical planners. In: AAAI. (2010)
5. Sohrabi, S., Riabov, A., Udrea, O.: Plan recognition as planning revisited. In: IJCAI. (2016)
6. Shvo, M., Sohrabi, S., McIlraith, S.A.: An ai planning-based approach to the multi-agent plan recognition problem (extended version). Technical Report CSRG-636, Department of Computer Science, University of Toronto (February 2018)
7. Gerevini, A., Saetti, A., Serina, I.: LPG-TD: a fully automated planner for PDDL 2.2 domains. In: ICAPS. (2004)
8. Nguyen, T.A., Do, M.B., Gerevini, A., Serina, I., Srivastava, B., Kambhampati, S.: Generating diverse plans to handle unknown and partially known user preferences. AIJ **190** (2012) 1–31