Active Goal Recognition

Maayan Shvo, Sheila A. McIlraith

Department of Computer Science, University of Toronto, Toronto, Canada

Vector Institute, Toronto, Canada

{maayanshvo, sheila}@cs.toronto.edu

Abstract

The objective of goal recognition is to infer a goal that accounts for the observed behavior of an actor. In this work, we introduce and formalize the notion of *active goal recognition* in which we endow the observer with agency to sense, reason, and act in the world with a view to enhancing and possibly expediting goal recognition, and/or to intervening in goal achievement. To this end, we present an algorithm for active goal recognition and a landmark-based approach to the elimination of hypothesized goals which leverages automated planning. Experiments demonstrate the merits of providing agency to the observer, and the effectiveness of our approach in potentially enhancing the observational power of the observer, as well as expediting and in some cases making possible the recognition of the actor's goal.

1   Introduction

Goal recognition is the problem of inferring a goal that accounts for the observed behavior of an actor. We observe that in many real-world settings the observer cannot passively wait for the natural evolution of the world in order to disambiguate the actor's goal. Indeed we advocate for the observer to be active and proactive in the goal recognition process.

To advance this view of goal recognition, we introduce and formalize the notion of *active goal recognition* where we endow the observer with agency to actively sense, act, or react as part of the goal recognition process. An active observer therefore has the potential to increase observational power, expedite, and in some cases make possible (where without observer agency, it was impossible) the recognition of the actor's goal. We cast the observer's task of deciding how to act in service of the recognition process as a contingent planning task, and discuss some of the challenges to this characterization in the general case.

Goal recognition is situated within the broader field of plan, activity, and intent recognition whose various sub-problems are concerned, together, with inferring not only the actor's 'top-level' goal, but also the plan that achieves this goal (where the plan being carried out by the actor is possibly independent of a specific goal), and even the 'low-level' actions, or activities of the actor (Sukthankar et al. 2014). While approaches to goal (and plan) recognition have

often matched between the observed behavior of the actor and pre-defined plan libraries (e.g., (Kautz and Allen 1986; Avrahami-Zilberbrand and Kaminka 2005)), the past decade has seen the rise of plan and goal recognition approaches that utilize planning domains as a form of generative model (e.g., (Ramírez and Geffner 2009; Sohrabi, Riabov, and Udrea 2016)). Further, a body of previous work has identified that performing the recognition *online*, while the actor's plan is still being carried out, is critical and has thus proposed efficient approaches geared towards online recognition (e.g., (Vered et al. 2018)). In the *offline* setting, the body of work on goal recognition design can also be seen to endow the observer with agency to modify the environment (Keren, Gal, and Karpas 2014), improve observability via sensor placement (Keren, Gal, and Karpas 2016) or to provide the actor with information (Keren et al. 2020) to benefit recognition during the eventual runtime. Most closely related to our work is a body of work that has provided the observer with agency to act in the world during runtime and disambiguate the hypothesis space. Kabanza et al. (2010) propose to provoke the actor into performing actions that will reveal its goal and Mirsky et al. (2018) formulate the sequential plan recognition process wherein the actor is queried about its ongoing plan, leading to hypothesis elimination. While our motivations partially overlap, these works require the observer to affect the state of the world or interact with the actor in order to eliminate hypotheses.

The main contributions of this paper are: (1) a formalization of active goal recognition; and (2) an algorithm for active goal recognition and a landmark-based approach for hypothesis elimination which leverage automated planning. We evaluate our approach in a diversity of domains using off-the-shelf planning and goal recognition tools. Our evaluation demonstrates the merits of providing agency to the observer, and the effectiveness of our approach in potentially enhancing the observational power of the observer.

2   Background

Automated planning   is the task of selecting a goal-leading plan based on a high-level description of the world. In this work we give agency to our observer to sense, act, or react. The determination of observer behavior fundamentally involves planning under partial observability with sensing actions. Such planning is typically referred to as *con-*

tingent planning (e.g., (Albore, Palacios, and Geffner 2009; Brafman and Shani 2012)). Contingent planning can be realized in either an offline or an online setting. In an offline setting, contingent planning yields a **contingent plan** that takes the form of a *conditional plan*, since the sensing action is not immediately executed to inform future action selection. In an online setting a contingent plan is conceived and the sensing immediately executed, updating the agent belief state and informing ongoing planning.

Following Bonet and Geffner (2011), we specify the partially observable planning problem with sensing actions (PPOS) in a STRIPS-like language, where actions can have conditional effects. A PPOS problem is a tuple $\langle \mathcal{F}, \mathcal{A}, \Omega, \mathcal{I}, G \rangle$ where $\mathcal{F}$ is the set of fluent atoms, $\mathcal{A}$ is the set of actions, Ω is the set of sensing actions, $\mathcal{I}$ is a set of clauses over $\mathcal{F}$ that determines the initial state, and G is a conjunction of atoms over $\mathcal{F}$ determining the goal condition. The specification is interpreted over a set of states, which are truth valuations to the atoms in $\mathcal{F}$. We say a literal l holds in a state s *iff* s assigns l to be true. Specifically, as $\mathcal{I}$ is a set of clauses, $\mathcal{I}$ would hold in a number of states; the *belief state* b is the set of all states where $\mathcal{I}$ holds. By extension, we say a formula ϕ holds in b *iff* ϕ holds in every state $s \in b$.

For $a \in \mathcal{A}$, the precondition of the action, $\text{PRE}(a)$, is a conjunction of atoms and the effect $\text{EFF}(a)$ is a set of pairs $\langle c, l \rangle$ that capture a's conditional effects. A sensing action $\omega \in \Omega$ has a precondition $\text{PRE}(\omega)$, which is a conjunction of fluent literals, and $\text{OBS}(\omega)$, which is the fluent literal that is observed by the sensing action. We say an action a is *applicable* in a state s *iff* $\text{PRE}(a)$ holds in s. We say a is applicable in a belief state b *iff* a is applicable in every $s \in b$. When performing a in b, a successor belief state b' is defined by performing a in each $s \in b$. When performing a sensing action ω in b, the successor belief state b' is the maximal set of states in b that agree on $\text{OBS}(\omega)$. By extension, a sequence $a_0, \ldots, a_k$, possibly involving sensing actions, is applicable in b if a_0 is applicable in b, resulting in a successor belief state b_1, and inductively, a_i is applicable in b_i, finally resulting in b_k. We say that a belief state b' is reachable from b if there is some sequence of actions (both sensing and non-sensing) that results in b' when applied to b.

A *conditional plan* (or policy) τ is induced by the outcomes of sensing actions and advises the next action to be taken based on the result of sensing actions (Geffner and Bonet 2013). We say that a conditional plan τ **solves** a PPOS problem $\langle \mathcal{F}, \mathcal{A}, \Omega, \mathcal{I}, G \rangle$ *iff* the executions advised by τ are applicable in the belief state b for $\mathcal{I}$, and result in belief states b^* where the goal condition G holds. We say that a PPOS problem $\langle \mathcal{F}, \mathcal{A}, \Omega, \mathcal{I}, G \rangle$ is a *classical* planning problem if $\mathcal{I}$ defines a complete initial state (i.e., $|b| = 1$ where b is the belief state for $\mathcal{I}$) and $\Omega = \emptyset$. We use the tuple $\langle \mathcal{F}, \mathcal{A}, \mathcal{I}, G \rangle$ to denote a classical planning problem. Note that solutions to a classical planning problem are simple sequences of actions (plans) since $\Omega = \emptyset$.

In this paper, we assume unit action costs and thus plan cost is equivalent to plan length, and optimal plans are the shortest plans. We use $\Pi^*(G)$ to denote the set of optimal plans for some goal G.

Goal recognition is the task of inferring a goal that accounts for the observed behavior of an actor.

Definition 1 (Goal Recognition Problem) *A goal recognition problem is a tuple* $\langle \Sigma, \mathcal{I}, \mathcal{G}, O \rangle$ *with* $\Sigma = \langle \mathcal{F}, \mathcal{A} \rangle$ *and where* $\mathcal{F}$ *and* $\mathcal{A}$ *are sets of fluents and actions, respectively,* $\mathcal{I}$ *is a set of clauses over* $\mathcal{F}$ *that determines the initial state,* $\mathcal{G} = \{G_1, ..., G_n\}$ *is a finite set of goals (henceforth hypotheses) where each* G_i *is a set of positive literals interpreted as a conjunctive formula, and* $O = o_1, ..., o_m$ *is a sequence of observations such that each observation* o_i *is a pair* (α_i, ϕ_i) *comprising an observed action,* $\alpha_i \in \mathcal{A}$, *and a set of literals,* ϕ_i *drawn from* $\mathcal{F}$.

The actor is assumed to be pursuing one and only one goal $G^* \in \mathcal{G}$. We emphasize the nonstandard definition of observations. o_i in an observation pair (α_i, ϕ_i) where α_i corresponds to an action drawn from $\mathcal{A}$ and ϕ_i corresponds to the truth or falsity of some properties of the state, drawn from $\mathcal{F}$, immediately following the execution of α_i, in the case where α_i is nonempty. For example, the observation *(move(lab,corridor),handsEmpty)* denotes that the actor was observed to perform the action *move(lab,corridor)* and the property *handsEmpty* was observed in the resulting state. 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 empty. The addition of state properties to the observations can be critical to goal recognition. Properties of state can influence the actor's decision making, and they can also provide clues to the occurrence of past unobserved actions.

Given a goal recognition problem, $\langle \Sigma, \mathcal{I}, \mathcal{G}, O \rangle$, a sequence of world-altering actions $a_1, \ldots, a_n$ **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 ϕ_j holds in b' which is the belief state that results from performing $a_1, \ldots, a_{f(j)}, j = 1, \ldots, m$, in b for $\mathcal{I}$. Following Ramírez and Geffner (2009), a **solution** to a goal recognition problem $\langle \Sigma, \mathcal{I}, \mathcal{G}, O \rangle$ is a set of hypotheses $\mathcal{G}' \subseteq \mathcal{G}$ such that for every goal $G \in \mathcal{G}'$ there exists some $\pi \in \Pi^*(G)$, the set of optimal plans for G, such that π *satisfies* O. We denote this set of hypotheses $\mathcal{G}^*$.

3 Active Goal Recognition

In standard accounts of goal recognition, the observer is a passive observer. In this section we introduce the notion of *active goal recognition* wherein we endow the observer with agency to actively sense, act, or react as part of the goal recognition process, thereby making the observer a first-class citizen in the recognition process. In tandem with the actor's pursuit of its goal, the observer executes a sequence of sensing and world altering actions. In the simplest case, an observer will simply perform a sequence of *sense* actions that evaluate the truth value of a fixed set of propositions. In more sophisticated cases, the observer will determine what to sense based on the candidate goals under consideration, interleaving world-altering actions (e.g., moving its location, picking up an object to look underneath it) with sensing in order to realize more purposeful observations or perhaps, as

we discuss briefly in Section 7, to go beyond goal recognition and actively intervene in the actor's pursuit of its goal. Such contingent plans—what we refer to here as *observer plans*—have the potential to increase observational power, expedite, and in some cases make possible (where without observer agency, it was impossible) the recognition of the actor's goal.

The definition of active goal recognition that follows allows for two views of the world—that of the actor and that of the observer, each with its own view of the initial state, and its own sets of actions, with the observer endowed with both world-altering and sensing actions. These actions dictate the subset of fluents that the actor (resp. observer) can affect, and/or observe. Typically the actor is assumed to be operating with complete information about the fluents related to the achievement of its goals.

Definition 2 *An **active goal recognition problem** is a tuple $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$ where $\Sigma = \langle \mathcal{F}, \mathcal{A}_A \cup \mathcal{A}_O \cup \Omega_O \rangle$, the fluents and the actions, comprising the world-altering actions of the actor and the world-altering and sensing actions of the observer; and the initial state $\mathcal{I} = \langle \mathcal{I}_A, \mathcal{I}_O \rangle$ comprising two sets of clauses over $\mathcal{F}$, describing the initial state of the actor and observer, respectively; and finally τ which is the contingent plan corresponding to the observer plan.*

Notation: For notational convenience, we use Σ_A to denote the pair $\langle \mathcal{F}, \mathcal{A}_A \rangle$, the fluents and actions of the actor.

Note that execution of the observer plan, τ, in whole or in part, yields a sequence of observations, $O = o_1, ..., o_m$, where each o_i is a pair (α_i, ϕ_i) per Definition 1. This in turn reduces the active goal recognition problem to a goal recognition problem.

We have postponed detailed discussion of the observer plan, τ, to Section 4. In some cases, the plan is determined a priori, and takes the form of a conditional plan. In other cases it is desirable to generate it on the fly. In this latter case, τ is initially empty and active goal recognition yields a second problem—that of generating the observer plan.

Definition 3 *A **solution** to an active goal recognition problem $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$ is a set of hypotheses $\mathcal{G}' \subseteq \mathcal{G}$ where for each hypothesis $G \in \mathcal{G}'$ there exists a plan $\pi \in \Pi^*(G)$ that satisfies the sequence of observations O resulting from the (partial) execution of the observer plan τ.*

Example 1. Consider a scenario (proposed by Pozanco et al. (2018)) where a terrorist (the actor) has committed an attack in the center of a city and is attempting to leave the city by either plane, train, or bus. The police (the active observer) are trying to prevent the actor from leaving the city and are using cameras located in various locations around the city in an attempt to locate the actor. Unfortunately, the police do not have sufficient resources to set up more than one blockade/control point along the various escape routes (to prevent the actor from escaping) and must therefore first recognize the actor's goal. To this end, in addition to the fixed cameras, the police have at their disposal a drone that they can deploy

to scan various locations in the city in an attempt to obtain additional observations pertaining to the actor's location. To illustrate, we partially model Example 1 as an active goal recognition problem.

$droneSense(\$loc) \in \Omega_O$

   PRE($droneSense$) := at(Drone,$\$loc$)
   OBS($droneSense$) := at(Terrorist,$\$loc$)

$cameraCityHallSense \in \Omega_O$

   PRE($cameraCityHallSense$) := on(CameraCityHall)
   OBS($droneSense$) := at(Terrorist,CityHall)

$\mathcal{G} = \{at$(Terrorist,Airport),at(Terrorist,TrainStn), at(Terrorist,BusStn)$\}$

At the outset, the observer has not started to execute τ, there are no observations, and thus $\mathcal{G}^* = \mathcal{G}$. The observer plan τ advises the observer to first perform *cameraCityHallSense* followed by *droneSense* where $\$loc$ is Zoo (which necessitates deploying the drone to the zoo to satisfy at(Drone,Zoo)), and upon executing these two actions the police observe whether or not at(Terrorist,CityHall) and at(Terrorist,Zoo) hold. Assume that at(Terrorist,Zoo) was observed, and further assume that the zoo is on the optimal path to the airport from the city center but not on any optimal path leading from the city center to either the bus station or the train station. Given the observations induced by this partial execution of τ, it follows that $\mathcal{G}^* = \{at$(Terrorist,Airport)$\}$.

4 The Observer Plan

The active goal recognition problem includes the observer plan, τ, which we identify as a contingent plan (Definition 2). In some cases, this contingent plan is determined a priori, and takes the form of a conditional plan. In other cases it is initially a null plan and is automatically generated on the fly.

Given the active goal recognition problem, $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$, the goal for τ is typically to generate a plan that results in a unique hypothesis, a $\mathcal{G}' \subseteq \mathcal{G}$ where $|\mathcal{G}'| = 1$. In other cases, goal recognition might want to eliminate a particular hypothesis from $\mathcal{G}$ (e.g., it might want to know that the terrorist is not heading to the airport). From the perspective of the observer, such goals are fundamentally *epistemic*, since they talk about the observer's state of knowledge, rather than the state of the world. For example, the goal might be to *know* (or believe) G_1 is being pursued by the actor and to know that G_2 and G_3 are not being pursued. This is of course different from knowing G_1 and knowing $\neg G_2$ and $\neg G_3$. We do not want to know that the goals *currently* hold (resp. do not hold), but rather that they are (resp. are not) being pursued. Further, since our observer is endowed with both world-altering actions as well as sensing actions, we don't want our observer to commandeer a helicopter to transport the terrorist to the airport and derive certainty by changing the state of the world.

The pursuit of epistemic goals has been studied most recently in the context of the burgeoning field of epistemic

planning (e.g., (Baral et al. 2017; Bolander 2017)) and another approach to addressing this class of problems would be to cast them as epistemic planning problems. Alternatively, we can exploit a contingent planner that accepts epistemic goals. Indeed in previous work Baier, Mombourquette, and McIlraith (2014) developed an approach to contingent planning with epistemic goals that obviates the need for true epistemic planning via a compilation method to contingent planning. Generation of our contingent plans with epistemic goals could be realized using such a planner, though the problem of distinguishing between knowledge of the pursuit of a goal vs knowledge of goal realization still remains.

For the purposes of this paper, we take a more pragmatic approach and in Section 5 propose a goal recognition system that generates and executes small observer plans for specific hypothesis-related goals, interleaving these with the updating of $\mathcal{G}'$, the candidate hypotheses (goals) that remain under consideration, via standard goal recognition systems. In the subsection that follows, we introduce theory central to the development of this system.

4.1 Testing Hypotheses

Once you eliminate the impossible, whatever remains, no matter how improbable, must be the truth.
- Arthur Conan Doyle

To understand how to achieve hypothesis-related goals relating to goal recognition we need to understand how acting and sensing can eliminate hypotheses from a candidate goal set $\mathcal{G}$, and we must identify any associated properties we might want our observer plans to have in order to realize the intended consequences of our observer plan τ. To this end, we revisit ideas first introduced in work by McIlraith and Reiter (1992) on hypothetical reasoning, adapting them to the context of goal recognition.

Most sensing actions simply return the truth or falsity of a fluent or set of fluents relative to a particular state. The definition of a test augments this fundamental notion of sensing with a set of conditions that must be true for the test to be realized. For example, if the observer wants to see whether the terrorist is at the zoo entrance, it must position a camera at the zoo entrance and then sense (i.e., look). The location of the camera at the appropriate location, together with any other preconditons of the sense action constitute the initial conditions for the test that determines if the terrorist is at the zoo entrance.

Definition 4 (Simple Test) *A **simple test** is a tuple $\langle C, \omega \rangle$ where C, the initial conditions for the test, is a conjunction of literals and $\omega \in \Omega$ is a sensing action.*

Tests may also be *complex*, comprising a series of simple tests interleaved with world altering actions whose execution satisfies the initial conditions necessary to perform the various simple tests (McIlraith and Reiter 1992). One would typically expect the observer plan τ to be a complex test (possibly) comprising multiple simple tests (whose outcomes serve as additional observations) interleaved with the actions the observer must perform in the world to realize the initial conditions for each simple tests.

Definition 5 (Test Outcome) *The **outcome** of a simple test $\langle C, \omega \rangle$ is $\text{OBS}(\omega)$.*

Definition 6 (Refutation) *Given an active goal recognition problem $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$, a test outcome μ **refutes** a hypothesis $G \in \mathcal{G}$ iff no plan for G exists that satisfies the concatenation of observation $(\emptyset, \mu)$ to observation sequence O, resulting from the execution of τ.*

In Example 1, a test outcome μ may inform the police that all roads leading to the airport are blocked. μ then refutes the hypothesis that the terrorist is attempting to escape via plane, i.e., *at(*Terrorist,Airport*)*.

Note that the definition of refutation is stronger than many current-day goal recognition criteria (including ours) that appeal to the set of *optimal* plans for membership of $G \in \mathcal{G}$.

Definition 7 (Confirmation) *Given an active goal recognition problem $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$, a test outcome μ **confirms** a hypothesis $G \in \mathcal{G}$ iff there exists a plan $\pi \in \Pi^*(G)$ that satisfies the concatenation of observation $(\emptyset, \mu)$ to observation sequence O, resulting from the execution of τ.*

In Example 1, if a test outcome μ places the terrorist on the optimal route between the city center and the airport, then μ confirms *at(*Terrorist,Airport*)*. A refuting or confirming test outcome may *eliminate* a hypothesis depending on the criteria by which a hypothesis space is defined (McIlraith and Reiter 1992). In the active goal recognition setting, we say that the set of hypotheses $\mathcal{G}$ has the following properties:

Proposition 1 *A test outcome μ **eliminates** all hypotheses $G \in \mathcal{G}$ that are refuted by μ.*

Proposition 2 *A test outcome μ **eliminates** all hypotheses $G \in \mathcal{G}$ that are not confirmed by μ.*

Domains and Observation Noise, and Suboptimality. Our definitions of refutation and confirmation (and thus of elimination) presuppose an accurate domain axiomatization, noise-free observations, and a rational actor that single-mindedly follows optimal paths to a goal. These assumptions (or slightly relaxed variants that tolerate paths that are 'more or less' optimal) are present in most previous work on cost-based goal and plan recognition (e.g., (Ramírez and Geffner 2009)). While reasonable in some setting these assumptions can prove too strong and may cause the recognizer's performance to suffer (Masters and Sardina 2019). To handle uncertainty, including stochasticity in the domain model and noise in observations, goal recognition systems can maintain a probability distribution over belief states and hypotheses.

Properties of Observer Plans. Observers are endowed with both sensing actions and world-altering actions that enable them to (at least) change the state of the world in order to realize observer plans in service of goal recognition. In Section 7, we discuss the case where observers may wish to aid or impede actor goals. That scenario notwithstanding, an important property for an observer plan τ is that it

be non-intervening. We say that τ is a **non-intervening observer plan** if for every hypothesis $G \in \mathcal{G}$, the set of plans π, whose execution achieves G are preserved under the execution of τ. That is, the execution of the observer plan does nothing that would preclude the actor from executing a plan to achieve a goal in $\mathcal{G}$. This is a strong assumption and may not always be desirable.

For example, the observer may introduce an obstacle along some path and test the actor's next action which, in the altered domain, will make clear the actor's goal (similarly to the *offline* goal recognition design setting (Keren, Gal, and Karpas 2014)). Such a plan would be considered an *intervening* observer plan, but it has clear utility in goal recognition.

Another potentially desirable property of observer plans is that obfuscation, privacy, and safety constraints may be enforced when the observer does not wish for its actions to be observable by the actor or for its goal to be inferred, or when it is desirable to preclude the observer from endangering itself or the actor. This can be important if the actor's awareness of the observer's actions or their consequences in some way influences the actor's realization of its goal. Such additional properties can be realized as additional (perhaps temporally extended) state constraints that are enforced within the plan generation process. More generally, we would like to enforce that the observer's actions do not influence the actor's decision making.

5 Computing the Observer Plan

In the previous section, we suggested various hypothesis-related goals pertaining to goal recognition. In this section, we define an algorithm for active goal recognition that generates and executes small observer plans for specific hypothesis-related goals, interleaving these with the updating of the set of hypotheses $\mathcal{G}'$.

Given an active goal recognition problem $P = \langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$, where τ is empty, Algorithm 1 returns the set of hypotheses $\mathcal{G}' = \mathcal{G}^* \subseteq \mathcal{G}$. In Line 1, the function GENERATEOBSPLAN takes P as input and generates the observer plan τ which is then executed in Line 2 by the function EXECUTEOBSERVERPLAN. In Line 3, RECOGNIZEGOAL is given as input a goal recognition problem $\langle \Sigma_A, \mathcal{I}_A, \mathcal{G}, O \rangle$ (where O results from the execution of τ) and returns the set of hypotheses $\mathcal{G}' = \mathcal{G}^* \subseteq \mathcal{G}$ as discussed in Section 2. We illustrate a run of Algorithm 1 at the end of this section.

Algorithm 1

Require: An active goal recognition problem
 $P = \langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$
1: $\tau \leftarrow$ GENERATEOBSPLAN(P)
2: $O \leftarrow$ EXECUTEOBSERVERPLAN(τ)
3: $\mathcal{G}' \leftarrow$ RECOGNIZEGOAL($\langle \Sigma_A, \mathcal{I}_A, \mathcal{G}, O \rangle$)
4: RETURN $\mathcal{G}'$

5.1 Landmark-based Hypothesis Elimination

We leverage the notions of tests and hypothesis elimination to realize the function GENERATEOBSPLAN which generates an observer plan τ given an active goal recognition problem $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$. We appeal to landmarks and focus on generating an observer plan τ on the fly comprising a single simple test (and the actions necessary to execute the test). We do so by constructing a PPOS problem whose solution is the observer plan τ. We leverage *fact landmarks* to construct the PPOS problem and guide test selection of observations. Following Hoffmann et al. (2004):

Definition 8 (Fact Landmark) *Given a classical planning task $P = \langle \mathcal{F}, \mathcal{A}, \mathcal{I}, G \rangle$, a formula l is a fact landmark in P iff l is true at some point along all plans that solve P.*

We do not consider disjunctive landmarks. Let $\mathcal{L}_G$ be the set of landmarks for some goal $G \in \mathcal{G}$ given a classical planning problem $\langle \mathcal{F}, \mathcal{A}, \mathcal{I}, G \rangle$. The observer's set of sensing actions, Ω_O, is assumed to be augmented with sensing actions $\omega \in \Omega_O$ that encapsulate simple tests $\langle C, \omega' \rangle$, where $\text{PRE}(\omega) = \text{PRE}(\omega') \wedge C$ and $\text{OBS}(\omega)$ is the test outcome μ. Recalling Definitions 6 and 7:

Proposition 3 *Given an active goal recognition problem $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$, a test outcome μ refutes a hypothesis $G \in \mathcal{G}$ if $\neg\mu$ is a landmark for G and no plan for G exists that satisfies the concatenation of observation $(\emptyset, \mu)$ to observation sequence O.*

Proposition 4 *Given an active goal recognition problem $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$, a test outcome μ confirms a hypothesis $G \in \mathcal{G}$ if μ is a landmark for G (i.e., $\mu \in \mathcal{L}_G$).*

Proposition 4 follows from Definition 7 since if μ is a landmark for G, it is made true at some point along all plans in $\Pi^*(G)$. Thus, there exists a prefix $a_1, \ldots, a_m$ of some plan in $\Pi^*(G)$ which 'satisfies' μ. Following Definition 7, a test outcome μ does *not* confirm a hypothesis $G \in \mathcal{G}$ *iff* there does not exist an optimal plan for G, π, such that μ holds following some prefix of π. Recall that, following Proposition 2, a hypothesis $G \in \mathcal{G}$ is eliminated if G is not confirmed by a test outcome μ. We weaken this definition to ease computation and make use of landmarks as a guide for the selection of tests and subsequent gathering of observations.

Proposition 5 *If a test outcome μ is a landmark for a hypothesis $G_i \in \mathcal{G}$ but not for a hypothesis $G_j \in \mathcal{G}$ (i.e., $\mu \in \mathcal{L}_{G_i}$ and $\mu \notin \mathcal{L}_{G_j}$), then μ weakly eliminates G_j.*

The most 'useful' landmarks to be tested by the observer are thus those that maximally (weakly) eliminate hypotheses in $\mathcal{G}$. To determine the 'usefulness' of a landmark, we extract (using Σ_A and $\mathcal{I}_A$) the set $\mathcal{L}_{\mathcal{G}}$ which comprises the sets of landmarks ($\mathcal{L}_G$) for each hypothesis $G \in \mathcal{G}$. We compute from $\mathcal{L}_{\mathcal{G}}$ the set of landmarks $\mathcal{L}_u$ with the highest *'landmark uniqueness value'* which represents the information value of a given landmark (Pereira, Oren, and Meneguzzi 2017, Equation 2, pg. 4). $\mathcal{L}_u$ comprises the 'most unique' landmarks such that each landmark $l \in \mathcal{L}_u$ belongs to a minimal number of $\mathcal{L}_G \in \mathcal{L}_{\mathcal{G}}$. For example, a landmark $l \in \mathcal{L}_u$ could be shared by two and only two hypotheses G_1 and G_2 and upon observing l, all hypotheses but G_1 and G_2 may be weakly eliminated (following Proposition 5).

Next, we augment $\mathcal{F}$ with a special predicate *done* and generate a set of actions $\mathcal{A}''$ and for each landmark $l \in \mathcal{L}_u$ we add an action a to $\mathcal{A}''$ where $\text{PRE}(a) = l$ and the effect of a is *done*. To ease exposition, we focus on hypothesis elimination via non-confirmation. However, our approach can be extended to address hypothesis elimination via refutation (following Proposition 1) by generating additional 'elimination' actions whose preconditions are landmarks with the property formalized in Proposition 3.

To ensure that the observer plan is non-intervening, we remove all actions $a \in \mathcal{A}_O$ where $\text{EFF}(a)$ includes a landmark (or a negated landmark) $l \in \mathcal{L}_G$ for any $G \in \mathcal{G}$. While this method of enforcing non-intervention is maximally prohibitive to the observer, it is minimally disruptive to the actor. We will explore other, less prohibitive methods of enforcing non-intervention in future work.

Finally, given an active goal recognition problem $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$, we construct a PPOS problem $R = \langle \mathcal{F}', \mathcal{A}'_O, \Omega_O, \mathcal{I}_O, G \rangle$, where $\mathcal{F}$ is augmented with *done*, $\mathcal{A}'_O = \mathcal{A}_O \cup \mathcal{A}''$, and $G = done$. R is then given to a planner that generates the observer plan τ that solves the planning task. As τ may include sensing actions, it may be the case that the results of these sensing actions will be such that G does not hold in some resulting belief states. In such a case, τ is considered a *weak* plan that reaches the goal under at least one possible set of action outcomes of the actions in the plan, as opposed to a *strong* plan which is a closed policy that achieves the goal in all resulting belief states (Cimatti et al. 2003). Consider Example 1 where τ may include the sensing action *droneSense*. The result of this action, which is the truth value of *at(Terrorist,$loc)*, may or may not satisfy the precondition of the relevant action in $\mathcal{A}''$ (i.e., that *at(Terrorist,$loc)* holds) and thus *done* may not hold following the execution of τ.

In Line 2, τ is executed and throughout its execution the state of the world is updated . We assume τ remains applicable throughout its execution and will address the ramifications of an ever-changing world that may render τ inapplicable in future work.

Theorem 1 (Complexity) *Given an active goal recognition problem $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$ and the corresponding PPOS problem used to generate the observer plan τ, $R = \langle \mathcal{F}', \mathcal{A}_O, \Omega_O, \mathcal{I}_O, G \rangle$, if $\mathcal{I}_O$ does not define a complete initial state and $\Omega_O \neq \emptyset$ then R is a conditional planning problem and deciding whether there exists a solution for R is 2-EXPTIME-complete.*

The proof follows straightforwardly from Rintanen's results on the complexity of conditional planning (Rintanen 2004).

Example 1 (continued). We return to our example and illustrate a run of Algorithm 1. In Line 1, we generate the observer plan using the landmark-based approach described previously in this section. Let *at(Terrorist,Library)* be a landmark in the set of most unique landmarks $\mathcal{L}_u$. *at(Terrorist,Library)* is a landmark for *at(Terrorist,Airport)* (but not for *at(Terrorist,TrainStn)* or *at(Terrorist,BusStn)*) and represents the fact that the terrorist is at the library which

is along the path from the city center to the Zoo, which, in turn, is located along the path from the city center to the airport (and is itself a landmark for *at(Terrorist,Airport)*). Thus, one of the newly generated elimination actions in $\mathcal{A}''$ has the precondition *at(Terrorist,Library)* and the effect *done*. The generated plan τ might therefore advise deploying the drone to the library and performing the sensing action *droneSense*. Further, assume that following the execution of τ in Line 2, the drone observes *at(Terrorist,Library)*. Thus, following Line 3, $\mathcal{G}' = \{at(Terrorist,Airport)\}$ (since there does not exist an optimal plan for *at(Terrorist,TrainStn)* or *at(Terrorist,BusStn)* that passes through through the library).

6   Experimental Evaluation

Recall that in standard accounts of goal recognition, the observer is a passive observer. In active goal recognition, we endow the observer with agency to actively sense, act, or react as part of the goal recognition process. An active observer therefore has the potential to increase observational power, expedite, and in some cases make possible (where without observer agency, it was impossible) the recognition of the actor's goal. In this section, we demonstrate the merits of an active observer with a set of experiments.

The objectives of our evaluation were four-fold: (1) to demonstrate that endowing the observer with agency to act in the world results in earlier recognition compared to a passive approach; (2) to illustrate that in some settings it is impossible for a passive observer to infer the actor's goal without agency to act in the world; (3) to illustrate that enforcing constraints on the observer's actions can mitigate disruptive behavior; and (4) to demonstrate the applicability of off-the-shelf conditional planners to active goal recognition. To this end, we constructed active goal recognition problems from a diversity of domains and ran them using an off-the-shelf conditional planner. We experimented with seven domains, with six being goal recognition benchmarks taken from an openly available repository based on the benchmarks developed by Ramírez and Geffner and later extended and published by Pereira and Meneguzzi (2017). The TERRORIST domain was obtained with thanks to the authors.

In INTRUSION DETECTION, first introduced by Geib and Goldman (2001), an attacker (the actor) may want to attack a set of servers (e.g., by vandalizing a server or stealing data from it). Actions available to the attacker include obtaining access to a server, as well as deleting, modifying, or downloading files and logs. The observer's actions include inspecting a server's status by accessing and inspecting various logs. In GRID, the actor moves in a grid with the goal of reaching a certain cell. Moving between some cells requires keys whose obtaining may require the actor to obtain additional keys and so on. The observer's actions include moving between cells and sensing whether certain keys have been taken by the actor. In LOGISTICS, the observer is trying to ascertain the destinations of various packages that are being transported by trucks and airplanes. The observer can inspect warehouse logs as well as communicate with control towers in various airports. Similarly, in DEPOTS and DWR the observer can inspect warehouse logs and in

ZENOTRAVEL the observer can communicate with various control towers. TERRORIST models Example 1.

We experiment with 15 problems for each domain. From each goal recognition problem, $\langle \Sigma, \mathcal{I}, \mathcal{G}, O_{full} \rangle$, we create a set $\mathcal{O}$ comprising $m + 1$ incrementally growing observation subsequences from $O_{full} = o_1, ..., o_m$ (i.e., $\langle \rangle, \langle o_1 \rangle$, $\langle o_1, o_2 \rangle, ..., \langle o_1, ..., o_m \rangle$). From each problem we also create m+1 active goal recognition problems $\langle \Sigma, \mathcal{I}, \mathcal{G}, \tau \rangle$ by: (1) augmenting Ω_O and $\mathcal{A}_O$ with actions that enable the observer to test landmarks; (2) constructing an incompletely specified initial state $\mathcal{I}_O$; and (3) constructing τ from the corresponding subsequence from $\mathcal{O}$ (i.e., every observation in the subsequence corresponds to an action in τ). We assume unit action cost and only experiment with traces that represent *optimal* plans executed by the actor.

Early Recognition. To simulate an online recognition setting, Algorithm 1 is run m+1 times for each goal recognition instance, with the m+1 derived active goal recognition problems. In line 1, in order to construct the PPOS problem R, we augment $\mathcal{A}_O$ with 'elimination actions', as described in Section 5. The landmarks are extracted using the landmark generator in the FAST DOWNWARD system (Helmert 2006) which extracts all landmarks (including the orderings between them) given a classical planning task. From the set of extracted landmarks, we only choose landmarks that hold as the result of some action $a \in \mathcal{A}_A$. In the future, the ordering between the landmarks could be used to inform the ordering of simple tests such that the first landmarks to be tested are those that are expected to be achieved first (i.e., should be achieved next along any plan that reaches the goal). To avoid calling the landmark generator every time we run Algorithm 1, we store the cached landmarks and maintain bookkeeping to determine which landmarks have been achieved. We enforce non-intervention by removing all actions $a \in \mathcal{A}_O$ where EFF(a) includes a landmark (or a negated landmark) $l \in \mathcal{L}_G$ for any $G \in \mathcal{G}$. Finally, we give the constructed PPOS problem to the conditional planner PO-PRP (Muise, Belle, and McIlraith 2014) which returns the observer plan τ (which is concatenated with the 'fixed' observation actions corresponding to the observations from O_{full}. Note that PO-PRP is an offline conditional planner and we will explore the use of online contingent planners such as CLG (Albore, Palacios, and Geffner 2009) in future work. In Line 2, we simulate τ's execution. If the landmark l being tested belongs to G^* (which is known to the system), we randomly select a test outcome μ (i.e., either a 'negative' outcome where l does not hold or a 'positive' one where l holds). Otherwise, the test outcome is negative.

In Line 3 we realize the function RECOGNIZEGOAL using two different recognizers. The first recognizer we use was proposed by Ramírez and Geffner (2009) (obtained from https://sites.google.com/site/prasplanning/ and henceforth called RG) which we run with the FAST DOWNWARD planner (Helmert 2006) coupled with an admissible heuristic. Since RG expects observations of actions, we replace the observations of test outcomes by the corresponding actions in $\mathcal{A}_A$. RG makes two calls to the planner each time it is called and is thus not geared towards an online goal recognition setting. The second recognizer we use is Vered et

Domain	$|\mathcal{G}|$	$|O_{full}|$	Active		Passive	
			T	CV	T	CV
INTRUSION	15	11.5	1.72	**79.4%**	**0.47**	66.3%
GRID	7.5	19.2	2.1	**51.7%**	**0.62**	40.2%
LOGISTICS	10	18.7	2.9	**61.4%**	**0.43**	48.3%
TERRORIST	6	7.3	4.2	**90.3%**	**1.3**	84.8%
DEPOTS	8.5	24.1	3.1	**52.4%**	**1.2**	37.8%
DWR	6.5	38.6	4.6	**65.9%**	**0.98**	44.1%
ZENOTRAVEL	7.5	13.3	3.8	**73.2%**	**1.27**	62.7%

Table 1: Comparison between an active approach (Lines 1-3 of Algorithm 1) and a passive one (Line 3 of Algorithm 1) in various domains using VERED. Each row describes averages over fifteen problems, where the columns stand for number of hypotheses ($|\mathcal{G}|$), total number of observations ($|O_{full}|$), average time in seconds to run the relevant part of Algorithm 1 (T), and convergence to the correct hypothesis (CV).

al.'s (2018) landmark-based online goal recognition system (henceforth called VERED). VERED implements Vered et al.'s 'Goal Mirroring with Landmarks' approach (Algorithm 4 in their paper). Rather than returning the set of hypotheses $\mathcal{G}^* \subseteq \mathcal{G}$ like RG, VERED returns a probability distribution over the set of hypotheses given the observations, $P(G|O)$. VERED efficiently utilizes landmarks to rank the likelihood of hypotheses in $\mathcal{G}$ as well as prune hypotheses that are deemed unlikely given the observed landmarks. These hypotheses are assigned a probability of 0 and removed from $\mathcal{G}$. VERED requires the set of extracted landmarks for the goal recognition instance and we therefore provide the algorithm with the cached landmarks so that the landmark generator need only be called once. Lastly, VERED calls a planner once and we use FAST DOWNWARD for this purpose. Since we use optimal traces, neither recognizer wrongly eliminates hypotheses. Finally, to compare between a passive approach and an active one we 'skip' Lines 1-2 of Algorithm 1 in the passive case and simply provide a goal recognition problem (with the current subsequence of observations) to the recognizer. Thus, in the passive case O is only updated with observations from O_{full}.

Early Recognition - Results. Following (Vered et al. 2018), we are interested in both efficiency and performance measures. For efficiency, we measure the average time T (in seconds) required for a run of the relevant part of Algorithm 1. I.e., Lines 1-3 for the active approach and Line 3 for the passive. For performance, Vered et al. define convergence to the correct answer (CV) as the percentage of *unrevealed* observations in O_{full} at the time when the recognizer converged to the correct hypothesis (or 0 if it failed to converge). Thus, higher values of CV are better as they indicate earlier convergence. Table 1 compares between an active setting and a passive setting and shows the average T and CV values across all problems and all 'time steps' for each domain. While the CV values in Table 1 indicate the average time step in which convergence was achieved, Figure 1 'breaks down' the convergence process and shows the average percentage of remaining (non-eliminated) hypotheses in $\mathcal{G} \setminus \{G^*\}$ as a function of the percentage of revealed observations in O_{full} across all GRID (left) and INTRUSION-DETECTION (right) problems. That is, when all
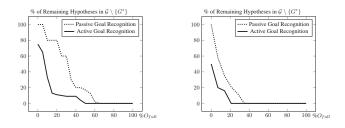
Figure 1: Average percentage of remaining hypotheses in $\mathcal{G} \setminus \{G^*\}$ as a function of the percentage of revealed observations in O_{full} across all GRID (left) and INTRUSION (right) problems. 0% signifies that all hypotheses but the actor's true goal G^* have been eliminated by the recognizer.

hypotheses but G^* have been eliminated (i.e., the actor's true goal is inferred and $\mathcal{G}' = \{G^*\}$ following Line 3), the set $\mathcal{G} \setminus \{G^*\}$ will be empty. The dotted and solid curves represent a passive setting and an active setting, respectively. For instance, in GRID, when 20% of O_{full} has been revealed, there were, on average, 10% remaining hypotheses in $\mathcal{G} \setminus \{G^*\}$ in the active setting compared to 80% in the passive setting. Results in Table 1 and Figure 1 pertain to VERED.

The T values comprise the average time it took to generate the observer plan in the active setting and a call to the recognizer in both active and passive settings. In all cases, PO-PRP generated the observer plan in under three seconds and FAST DOWNWARD extracted a set of landmarks for a given planning task in under two seconds. As a call to RG involves two calls to FAST DOWNWARD, the total runtime of Lines 1-3 when using RG was dominated by the call to RG with runtimes of tens (and sometimes hundreds) of seconds for the larger problems. We thus used RG as a baseline and do not report on these results in Table 1. In contrast, when using VERED the runtime dropped dramatically, as reflected in the low T values in Table 1.

With respect to our **first objective**, the results demonstrate that by endowing the observer with agency to act in the world, the observer can more expeditiously eliminate possible hypotheses (compared to the passive case where observations are simply 'given' to it), thus facilitating earlier recognition of the actor's goal (as measured by the CV values). Earlier recognition, in turn, can facilitate a better reaction on behalf of the observer. Note that, as will be discussed in Section 7, the observer will not necessarily achieve earlier recognition if the resources required for plan generation and execution (e.g., the time it would take to execute the observer plan) are not taken into consideration, as is the case in our experimentation.

Observational Power. Consider a scenario where transmission from the various sensors has suddenly stopped in the GRID domain and the observer is thus 'in the dark', with no further observations about the actor's movements at its avail. In the LOGISTICS domain, the observer may lose communication with the control tower and in TERRORIST the cameras may stop working due to a power failure or cyber attack. We simulate these scenarios in the three aforementioned domains by only updating O with observations from

O_{full} (and only executing the fixed observation actions in τ in the active case) until a random point where RG returns $|\mathcal{G}^*| > 1$ (i.e., when some hypotheses are not discriminable given the observations). We run Algorithm 1 as previously described and compare between a passive observer and an active one.

Unsurprisingly, in the passive case $\mathcal{G} \setminus \{G^*\}$ does not change after no additional observations from O_{full} are given to the passive observer. In the active case the observer generates a plan τ in every iteration of Algorithm 1 and eventually converges to the actor's true goal G^*. Wrt our **second objective**, the results demonstrate that in some settings a passive observer *cannot* discriminate between hypotheses without acting in the world, while an active observer may increase its observational power by acting in the world and obtaining additional observations.

Disruptive Behavior. We augment $\mathcal{A}_O$ with actions whose effects comprise of the predicate *done* and of negated landmarks which can render the goals in $\mathcal{G}$ unachievable. In GRID, the observer can change the locks on doors and steal keys; in INTRUSION, the observer can shut down servers and delete files; and in LOGISTICS, the observer can break a vehicles. We experiment by running Algorithm 1 with and without the enforcement of non-intervention via elimination of potentially disruptive actions.

Without enforcing non-intervention, PO-PRP (if possible) computes strong plans that refute hypotheses by causing goals to become unachievable rather than computing weak plans comprising sensing actions whose outcomes may or may not cause *done* to hold. With respect to our **third objective**, the results demonstrate that without enforcing constraints on the observer's actions, the latter may act in unintended (and possibly undesirable) ways. For instance, while it may not be the observer's goal to hinder the actor by rendering some hypotheses in $\mathcal{G}$ unachievable (as is the case in the counterplanning literature), the disruptive nature of its unconstrained plans may inadvertently do so.

Finally, with respect to our **fourth objective**, the results in this section demonstrate that an off-the-shelf conditional planner can be readily used to solve the constructed PPOS problems. However, our promising results do not come without caveats. The conditional planning tasks given to PO-PRP are rather simple, which allowed the planner to generate the policy in under two seconds. However, due to the high complexity of conditional planning (Rintanen 2004) the runtime can grow significantly as the problem size grows (as seen in Muise, Belle, and McIlraith's (2014) Table 1).

7 Beyond Recognition: Aiding or Impeding

To this point, we have focused on the task of active goal recognition—giving agency to the observer in service of recognizing the goal of the actor. Nevertheless, in most practical applications, goal recognition is not an end in itself. Indeed, it is typically the case that goal recognition is performed so that an agent (here, our observer) can perhaps aid or impede the actor's goal, as illustrated in our example.

The idea of endowing an observer with agency to aid or impede the actor has been proposed by a number of researchers in the context of goal or plan recognition (e.g.,

(Geib et al. 2016)). For example, Freedman and Zilberstein (2017) propose a method of assisting an actor by adopting the latter's presumed goal and generating a plan to achieve it. When the observer wishes to impede goal achievement, the body of work on counterplanning (e.g., (Freedman and Zilberstein 2017; Pozanco et al. 2018; Porteous and Lindsay 2019)) formulates a *counterplanning* task, wherein the observer (called an observing agent by Freedman and Zilberstein, a preventing agent by Pozanco et al., and a protagonist by Porteous and Lindsay) attempts to prevent the actor from achieving its presumed goal. Since these works use a goal recognition algorithm prior to assisting or impeding goal achievement, the algorithm proposed in Section 5 and demonstrated in our experiments can be straightforwardly used to augment these works as a black box that returns a subset of the hypotheses in $\mathcal{G}$.

Example 1 (continued). Recall that after observing *at(*Terrorist,*Library)*, $\mathcal{G}' = \{at(\text{Terrorist,Airport})\}$ following Line 3 of Algorithm 1. At this point, we can employ Pozanco et al.'s algorithm for domain-independent counterplanning (Pozanco et al. 2018, Algorithm 1, pg. 5) to generate a plan advising the observer (the preventing agent in their work) to set the blockade at the zoo (a landmark for *at(*Terrorist,*Airport)*), thereby stopping the terrorist.

More generally, and perhaps more importantly, endowing the observer with the ability to generate a contingent plan, τ, that achieves both ontic and epistemic goals simultaneously, as mentioned briefly in Section 4, would in principle enable the observer to eliminate sufficient hypotheses from $\mathcal{G}$ to generate a helpful plan. Indeed, achievement of different goals can often share subgoals, or even complete plans, so determination of a unique hypothesis need not be a precursor to an observer selecting and performing actions that aid or impede the actor. There is always a tension between acting in service of further refinement of $\mathcal{G}$ or in service of realizing or impeding those goals. Some of those trade-offs can be realized by a contingent planner that performs both sensing and acting and that is designed to minimize the overall cost of its plan. In such a setting, the cost of actions must be carefully considered, potentially including elements such as time, money, or actor frustration. Exploration of these ideas is left to future work.

8 Discussion and Summary

In this work, we have introduced and formalized the notion of active goal recognition in which we endow the observer with agency to sense, reason, and act in the world with a view to enhancing and possibly expediting goal recognition, and/or to intervening in goal achievement. We presented an algorithm for active goal recognition and a landmark-based approach that leverage automated planning to facilitate hypothesis elimination. Finally, we evaluated our approach on a diversity of domains and demonstrated the merits of an active observer compared to a passive observer.

As discussed in Section 1, our work is situated within the broader field of plan, activity, and intent recognition. Most closely related to our work is a body of work on plan and goal recognition that endows the observer with

agency to act in the world (and possibly interact with the actor) in order to disambiguate the hypothesis space *during runtime* (e.g., (Bisson et al. 2011; Mirsky et al. 2018)) as well as in an *offline* setting (Keren, Gal, and Karpas 2014; Keren et al. 2020). While Kabanza et al. (2010) independently propose to allow the observer to actively make observations, they use plan libraries and have not, to the best of our knowledge, implemented their approach. In all of these works, the observer selects tests whose outcomes (e.g., the actor's next action or response to a query) facilitate hypothesis elimination and are therefore important modalities at the avail of an active observer.

Our work is also relevant to the field of human-aware AI which seeks to build AI agents whose behavior is interpretable to the human (cast as the observer) in the loop. Recently, there has been notable interest in this field, evident by a large body of work with applications to human-human, human-machine, and machine-machine interaction. Chakraborti et al. (2019) present a comprehensive survey of the research landscape, as well as independently discuss the notion of an active observer and how the observer's agency should affect the decision making of an actor keen on obfuscating its plan or making it more legible. The ideas we have formalized in this work can be leveraged to facilitate an active observer in this context.

While in our experiments we assume a rational actor and noise-free observations, our approach can address potentially noisy observations by using outcomes of tests (i.e., additional observations) as evidence by which to weight hypotheses in a probabilistic setting. Finally, if probabilities are associated with hypotheses in $\mathcal{G}$, the observer can focus on hypotheses that will maximize information gain wrt $\mathcal{G}$. This may be desirable in general to minimize testing time on improbable hypotheses, as shown by Mirsky et al. (2018).

References

Albore, A.; Palacios, H.; and Geffner, H. 2009. A Translation-based Approach to Contingent Planning. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*.

Avrahami-Zilberbrand, D., and Kaminka, G. A. 2005. Fast and Complete Symbolic Plan Recognition. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI)*.

Baier, J. A.; Mombourquette, B.; and McIlraith, S. A. 2014. Diagnostic Problem Solving via Planning with Ontic and Epistemic Goals. In *Proceedings of the 14th International Conference on the Principles of Knowledge Representation and Reasoning (KR)*.

Baral, C.; Bolander, T.; van Ditmarsch, H.; and McIlrath, S. 2017. Epistemic Planning (Dagstuhl Seminar 17231). In *Dagstuhl Reports*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Bisson, F.; Kabanza, F.; Benaskeur, A. R.; and Irandoust, H. 2011. Provoking Opponents to Facilitate the Recognition of their Intentions. In *Proceedings of the 25th AAAI Conference on Artificial Intelligence (AAAI)*.

Bolander, T. 2017. A Gentle Introduction to Epistemic Planning: The DEL Approach. *arXiv preprint arXiv:1703.02192*.

Bonet, B., and Geffner, H. 2011. Planning under Partial Observability by Classical Replanning: Theory and Experiments. In *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI)*, 1936–1941.

Brafman, R., and Shani, G. 2012. A Multi-Path Compilation Approach to Contingent Planning. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI)*.

Chakraborti, T.; Kulkarni, A.; Sreedharan, S.; Smith, D. E.; and Kambhampati, S. 2019. Explicability? Legibility? Predictability? Transparency? Privacy? Security? The Emerging Landscape of Interpretable Agent Behavior. In *Proceedings of the 29th International Conference on Automated Planning and Scheduling (ICAPS)*.

Cimatti, A.; Pistore, M.; Roveri, M.; and Traverso, P. 2003. Weak, Strong, and Strong Cyclic Planning via Symbolic Model checking. *Artificial Intelligence* 147(1-2):35–84.

Freedman, R. G., and Zilberstein, S. 2017. Integration of Planning with Recognition for Responsive Interaction Using Classical Planners. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, 4581–4588.

Geffner, H., and Bonet, B. 2013. A Concise Introduction to Models and Methods for Automated Planning. *Synthesis Lectures on Artificial Intelligence and Machine Learning* 8(1):1–141.

Geib, C. W., and Goldman, R. P. 2001. Plan Recognition in Intrusion Detection Systems. In *Proceedings of the DARPA Information Survivability Conference and Exposition II (DISCEX'01)*, 46–55.

Geib, C.; Weerasinghe, J.; Matskevich, S.; Kantharaju, P.; Craenen, B.; and Petrick, R. P. 2016. Building Helpful Virtual Agents using Plan Recognition and Planning. In *Proceedings of the Twelfth Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE)*.

Helmert, M. 2006. The Fast Downward Planning System. *Journal of Artificial Intelligence Research* 26:191–246.

Hoffmann, J.; Porteous, J.; and Sebastia, L. 2004. Ordered Landmarks in Planning. *Journal of Artificial Intelligence Research* 22:215–278.

Kabanza, F.; Bellefeuille, P.; Bisson, F.; Benaskeur, A. R.; and Irandoust, H. 2010. Opponent Behaviour Recognition for Real-Time Strategy Games. In *Workshops at the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI)*.

Kautz, H. A., and Allen, J. F. 1986. Generalized Plan Recognition. In *Proceedings of the 5th National Conference on Artificial Intelligence (AAAI)*.

Keren, S.; Xu, H.; Kwapong, K.; Parkes, D.; and Grosz, B. 2020. Information Shaping for Enhanced Goal Recognition of Partially-Informed Agents. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI)*.

Keren, S.; Gal, A.; and Karpas, E. 2014. Goal Recognition Design. In *Proceedings of the 24th International Conference on Automated Planning and Scheduling (ICAPS)*.

Keren, S.; Gal, A.; and Karpas, E. 2016. Goal Recognition Design with Non-Observable Actions. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI)*.

Masters, P., and Sardina, S. 2019. Goal Recognition for Rational and Irrational Agents. In *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*.

McIlraith, S., and Reiter, R. 1992. On Tests for Hypothetical Reasoning. *Readings in model-based diagnosis* 89–96.

Mirsky, R.; Stern, R.; Gal, K.; and Kalech, M. 2018. Sequential plan recognition: An iterative approach to disambiguating between hypotheses. *Artificial Intelligence* 260:51–73.

Muise, C.; Belle, V.; and McIlraith, S. A. 2014. Computing Contingent Plans via Fully Observable Non-Deterministic Planning. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI)*.

Pereira, R. F., and Meneguzzi, F. 2017. Goal and Plan Recognition Datasets using Classical Planning Domains. *Zenodo*. https://doi.org/10.5281/zenodo.825878.

Pereira, R. F.; Oren, N.; and Meneguzzi, F. 2017. Landmark-based Heuristics for Goal Recognition. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*.

Porteous, J., and Lindsay, A. 2019. Protagonist vs Antagonist PROVANT: Narrative Generation as Counter Planning. In *Proceedings of the 18th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 1069–1077.

Pozanco, A.; E-Martín, Y.; Fernández, S.; and Borrajo, D. 2018. Counterplanning using Goal Recognition and Landmarks. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, 4808–4814.

Ramírez, M., and Geffner, H. 2009. Plan Recognition as Planning. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI)*.

Rintanen, J. 2004. Complexity of Planning with Partial Observability. In *Proceedings of the 14th International Conference on Automated Planning and Scheduling (ICAPS)*, 345–354.

Sohrabi, S.; Riabov, A. V.; and Udrea, O. 2016. Plan Recognition as Planning Revisited. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)*, 3258–3264.

Sukthankar, G.; Geib, C.; Bui, H. H.; Pynadath, D.; and Goldman, R. P. 2014. *Plan, Activity, and Intent Recognition: Theory and Practice*. Newnes.

Vered, M.; Pereira, R. F.; Magnaguagno, M. C.; Kaminka, G. A.; and Meneguzzi, F. 2018. Towards Online Goal Recognition Combining Goal Mirroring and Landmarks. In *Proceedings of the 17th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2112–2114.