Probabilistic Plan Legibility with Off-the-shelf Planners

Michele Persiani, Thomas Hellström
Umeå University, Umeå, Sweden
michelep@cs.umu.se,thomas.hellstrom@umu.se

Abstract

Legible planning is the creation of plans that best disambiguate their goals from a set of other candidates from an observer's perspective. In this paper we propose a method for legible planning for arbitrary PDDL domains, by extending previous research on legibility to classical planning without requiring to construct ad-hoc planners. We also discuss how the observer perspective may be estimated through a second order theory of mind that connects the planner's and the observer's task spaces. Our solution can for example be deployed in human-robot teaming scenarios, where an autonomous robot in a team can implicitly communicate its goal by producing legible plans. We present benchmark results on several PDDL planning domains. Our results generally show that plan legibility is a trade-off with plan efficiency, however, not all planning domains allows to increase legibility in the same way and a regularizing factor to balance legibility and efficiency was proved necessary.

Introduction

A main effort in recent research in Artificial Intelligence is to provide intelligent decision-making algorithms the ability to produce explanations, and to make their decisions understandable. Three aspects have been identified as crucial to allow understandability: their ability to produce trust, to allow interaction with their decisions, and to be transparent in their decision making processes (Fox, Long, and Magazzeni 2017). Even though such guidelines for *Explainable AI* are mostly tailored for recommender systems, they can be applied to most AI systems, and especially for algorithms producing plans of actions, for example to execute on an embodied agent such as a robot. In fact, a planner (robot) collaborating with humans should indeed propose useful plans, allow to mediate such plans, and be transparent in its actions. The concept *understandable robots* (Hellström and Bensch 2018) is often used in this context.

In this paper we focus on the *transparency* part of Explainable AI for planning algorithms, which broadly translates into the field of interpretable planning, i.e. the generation of plans that are understandable from a human point of view (Chakraborti et al. 2017; 2019; MacNally et al. 2018; Kulkarni, Srivastava, and Kambhampati 2019). Planning while keeping a human observer perspective into consideration is desired in most human-robot interaction, where the collaboration with a robot may become difficult whenever its observed actions and behaviors are not directly understandable by the human collaborator (Chakraborti et al. 2019).

A robot's behavior can be hard to understand for reasons such as model discrepancies (Chakraborti et al. 2017), where the robot uses a task model different than the human's, or an asymmetry of information between human and robot, e.g. differences in their beliefs. If not dealt with, such divergencies may lead to decreased interpretability of the robot's behavior, ultimately leading to a loss of trust in the robotic system (Hellström and Bensch 2018).

Interpretability in planning has been addressed using several terms with subtle yet relevant differences, such as *explicable* (Kulkarni et al. 2019; Zhang et al. 2017), *predictable* (Zhang et al. 2017) and *legible* planning (Dragan, Lee, and Srinivasa 2013; MacNally et al. 2018). In this paper we use legibility, which focuses on reducing ambiguity over other possible goals (Chakraborti et al. 2019). Legible planning requires definition of optimality criteria for plans that are not exclusively based on their execution cost, but rather on how the plan's goal can be successfully identified. In general, this requires to explicitly model the observer as an agent performing inference over the planner's actions and weighting the observations towards a set of possible goals.

The contributions in this paper extend the formalism for plan legibility (especially (Dragan, Lee, and Srinivasa 2013; Dragan and Srinivasa 2014)) to classical planning methods such as PDDL (McDermott 1998) without requiring the construction of ad-hoc planners. We further provide discussion on how legible planning involves the creation of plans that are legible inside the model used by the observer to evaluate goals, hence requiring to take into account the differences between this model and the task model being internally utilized by the planner. This difference between models has been addressed in adjacent topics such as model reconciliation (Chakraborti et al. 2017), but has, to the best of our knowledge, never been thoroughly put side-to-side with interpretable planning methods where planner and observer either share the same task model (e.g. trajectories in the Cartesian space) or have a different observation space. We argue that legible planning requires a second order theory of mind, which allows the planner to evaluate its own plans from a perspective of an observer that uses a possibly inaccurate mental model of the planner. Theory of mind acts as

the glue that connects their two task spaces.

We further provide an algorithm to produce legible plans using off-the-shelf PDDL planners, and measure its performance over several standard domains. Since PDDL is easy to use and fairly well known by the planning and robotics communities, an algorithm for legible planning based on PDDL can easily find practical implementations for robotic scenarios, in fields such as human-robot interaction and human-robot teaming, or for algorithms exploring theory of mind in the domain of planning.

The rest of the paper is structured as follows. In Section 2 we provide some brief relevant background on legible planning. In Section 3 we introduce a working definition of plan legibility in a probabilistic setting. Theory of mind is contextualized as a function which transforms the planner domain to the observer's perspective. In Section 4 we propose a procedure to compute legible plans using off-the-shelf PDDL planners. In Section 5 and 6 we provide an illustrative example and performance measures over a set of planning domains. In Section 7 we provide some conclusive remarks.

Background

Plan legibility for robotic manipulators was investigated in (Dragan, Lee, and Srinivasa 2013; Dragan and Srinivasa 2013; 2014). The authors proposed trajectories that are skewed in a way that avoids ambiguity of goal locations. The increased legibility of these trajectories was opposed to optimality in terms of travelled distance. The observer and the robot were respectively evaluating and producing legible plans in the Cartesian space of the manipulator. The same authors investigated how a manipulator's legibility affects human-robot interaction in a task where experimenters are asked to act according to the robot's predicted goal (Dragan et al. 2015). Their results showed that legible trajectories positively affected the perceived interpretability of the robot, increased interaction fluency, and overall induced a greater sense of collaboration during the interaction. They further showed, through a user study, that there was a significant preference for legible plans rather than functional trajectories (avoiding obstacles), or predictable ones (with minimal cost).

In (MacNally et al. 2018), the authors use POMDPs to model how the beliefs of an observer, in particular inferred goals, are affected by different action selection policies. Since sequences of actions performed by an actor (e.g. a robot) leads to inferred goal distributions in an observer's beliefs, the authors discuss how actions can be used to communicate what the goal being pursued is, hence allowing an *implicit* communication of the actor's goal that was highlighted as relevant whenever no explicit communication mechanisms between actor and observer exist. This implicit communication of the goal by acting in a way recognizable by observers was named transparent planning, which is equivalent to the notion of legible planning used by other authors. Legibility as an implicit communication mechanism is also explored in (Miura and Zilberstein 2020). Through the utilization of *Legible Markov Decision Processes* (MDPs) the authors show how it is possible to integrate an observer's beliefs into the actor's beliefs. An important point discussed by the authors is that legibility can be arbitrarily increased through a trade-off with planning cost, and a balance between the two is almost always required. In both (MacNally et al. 2018) and (Miura and Zilberstein 2020) there is no disambiguation between the planner's and the observer's task space, but rather they use the same underlying MDP.

Plan legibility is similarly discussed as a planning problem with controlled observability. By controlling how observations are made, an observer can be made to recognize the correct goal. Adversarial and cooperative observers could be respectively misled or informed. For example, (Kulkarni, Srivastava, and Kambhampati 2019) propose a controlled observability planning algorithm that can create plans that are more legible or obfuscating depending on the employed heuristic to search the state space. The authors propose for every goal to have a heuristic attracting the search, and by averaging these heuristics with appropriate weights they obtain plans that are more legible. The same authors extend this line of research by allowing for both cooperative and adversarial observers to be simultaneously present (Kulkarni, Srivastava, and Kambhampati 2020). In both papers, planner and observer have different task spaces where the ground truth observations, found by an observation model, are respectively mapped.

Probabilistic Plan Legibility

We make the assumption that discrimination of goals is performed by an observer while observing the planner performing a task. Hence, the legibility of a plan is based on the observations that the plans lead to and prior beliefs over the task, e.g. the state of the task prior to the attempt of recognizing its goal. Therefore, legible planning should produce observations in the observer perspective that, together with prior beliefs, lead to the best discrimination of the plan's true goal.

Taking a probabilistic approach, at any moment there is a set $G = \{\hat{g}, g_0, g_1, ..., g_n\}$ of possible goals the planner can aspire to. To be legible, a plan must produce observations that make its true goal $\hat{g}$ easily discernible from the other goals in G. Therefore, generating legible plans is equivalent to finding the plan that minimizes the difference between the true goal distribution that the planner has, and a goal distribution that the observer infers from from the observations. In this setting, the observer can be modeled as a probabilistic model for goal recognition $P(G|\Pi)P(\Pi)$, with Π being the set of possible plans, $P(\Pi)$ the prior probability distribution of observing sequences of actions (plans) while $P(G|\Pi)$ is the distribution of the goals given observed plans. Hence, the observer is modeled as estimating the planner's goal distribution by observing its sequences of actions. For a goal $g \in G$ we define the legibility of a plan π in this plan space as:

$$\text{legibility}(\pi, g) = \mathcal{H}(P_g(G), P(G|\pi)P(\pi)), \qquad (1)$$

where $\mathcal{H}$ is a similarity function of two probability distributions. $P_g(G)$ is the goal probability distribution for the goal,

with $P(G = g) = 1$ and $P(G) = 0$ for all other possible goals.

The most legible plan is denoted π_{legible} and defined as follows:

$$\pi_{\text{legible}} = \underset{\pi \in \Pi_{\hat{g}}}{\text{argmax}}\, \text{legibility}(\pi, \hat{g}), \tag{2}$$

where we consider only the plans $\Pi_{\hat{g}}$ that achieve the true goal $\hat{g}$ from a given starting condition I which is the same for every goal in G. Whether $\hat{g}$ is effectively best discerned from the other candidate goals is not assured, as it in some cases may be impossible to generate sufficiently legible plans due to constraints in the task space, or it might even be impossible to generate a plan at all.

Since Eq. 1 considers only complete plans, we further introduce the notion of n-legibility, indicating the legibility of a plan after n steps:

$$n\text{-legibility}(\pi, g) = \text{legibility}(\pi_{1..n}, g) \tag{3}$$

where $\pi_{1..n}$ is the plan comprising the n first steps of π.

We finally propose a generalized measure $\hat{n}$-legibility as a weighted average of the legibility of all of the plan prefixes $\pi_1, \pi_{1..2}, ..., \pi_{1..n}$:

$$\hat{n}\text{-legibility}(\pi, g) = \sum_{i \in 1..n} w_i \cdot \text{legibility}(\pi_{1..i}, g) = \sum_{i \in 1..n} w_i \cdot \mathcal{H}(P_g(G), P(G|\pi_{1..i})P(\pi_{1..i})), \tag{4}$$

$$\sum_i w_i = 1.$$

From this definition, the legibility of any step j can be obtained by setting $w_j = 1$. Other customized averages may, for example, give a greater importance of legibility during the early steps of plans.

Theory of Mind

Legible planning requires the planner to estimate the probability distribution over the possible goals, computed from the observer's perspective. While in previous discussion we implicitly assumed that both planner and observer share the same observation model, we now remove this assumption by introducing a second-order theory of mind (Meijering et al. 2011; Devin and Alami 2016).

Theory of mind relates to the ability of agents to attribute mental states and beliefs to themselves or other agents, and of creating a point of view of a situation in terms of beliefs, goals and intentions that is different from their own but rather belonging others. A first order theory of mind is expressed in the sentence "Bob thinks that Alice thinks X", or in other words Bob has an estimate of Alice's mental state, believing she's thinking X. Higher order theories deepen these levels of reasoning by extending the thinking chain. A second order reasoning would be "Carl thinks that [Bob thinks that Alice thinks X]"—with parenthesis added to highlight the recursion. In this case Carl holds an estimate of Bob's mental state. Arbitrary higher orders of reasoning follow the same incremental structure.

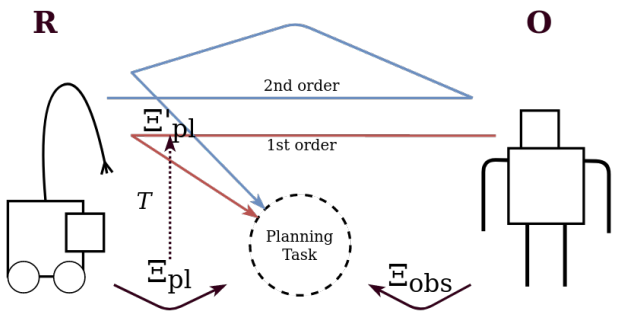

Figure 1: Legible planning requires a second order theory of mind, by which the planner, working e.g. on a mobile robot R, estimates how the observer O interprets the robot's actions by using a first order theory of mind of the robot. Legibility of plans is computed inside Ξ'_{pl} that is O's model of R.

In this paper we propose that legible planning requires a second order theory of mind as a model of how the observer infers the planner's goal. Consider the statement: "*O* thinks that *R*'s goal is *G*", where *O* is the observer, *R* is the robot, *G* is the robot's goal. The statement describes O's first-order theory of mind. At this level the belief about *R*'s goal belongs to *O*, and is not accessible by *R*. A second order chain of reasoning can be described as "*R* thinks that [*O* thinks that *R*'s goal is *G*]", where at the first level O uses her first order theory of mind, while at the second level the planner estimates the result of this inference by using its second order theory of mind. For R's legible planning, the second order reasoning can be reformulated as "*R*'s goal is that [*O* thinks that *R*'s goal is *G*]". The planner's action are legible not in the observer's task model (as was also proposed in previous literature), but rather inside the model that the observer uses to evaluate the planner's actions. This model, when implemented by *R*, effectively allows it to access *O*'s beliefs about him.

Following this reasoning in the setting of planning domains we define a second order theory of mind as a function T which allows to transform the plan instances $\boldsymbol{\Xi}_{\mathbf{pl}}$ utilized by the planner to the observer perspective of the same $\boldsymbol{\Xi}'_{\mathbf{pl}}$:

$$T = T_{pl} \circ T_{obs} : \boldsymbol{\Xi}_{\mathbf{pl}} \rightarrow \boldsymbol{\Xi}_{\mathbf{obs}} \rightarrow \boldsymbol{\Xi}'_{\mathbf{pl}} \tag{5}$$

such that for a specific instance Ξ_{pl} we can get a corresponding instance as $T(\Xi_{pl}) = \Xi'_{pl}$. T can also be seen as the composition of the first order theories of mind T_{pl} and T_{obs}, through which the planner makes an estimate of the observer's first order model, hence obtaining a second order model. This procedure is illustrated in Figure 1.

In PDDL, T could transform domains from $\boldsymbol{\Xi}_{\mathbf{pl}}$ by modifying its operators and predicates with corresponding operators and predicates taken from $\boldsymbol{\Xi}'_{\mathbf{pl}}$, and an associated problem instance by describing the initial and goal conditions in terms of truth values from $\boldsymbol{\Xi}'_{\mathbf{pl}}$. In this setting, plans produced in $\boldsymbol{\Xi}_{\mathbf{pl}}$ induce corresponding plans in $\boldsymbol{\Xi}'_{\mathbf{pl}}$.

To compute legibility the domains $\boldsymbol{\Xi}'_{\mathbf{pl}}$ are utilized, as they model how the observer perceives and integrate in its

beliefs the produced observations in the form of a plan. Notice that this model is possessed locally by the planner by its second order theory of mind, and once implemented does not require additional external input to execute a task legibly. Nevertheless, even after deployment, if $\mathbf{\Xi_{pl}}$ and $\mathbf{\Xi'_{pl}}$ result being sufficiently different from one another the observer can judge the plans as non legible. In these cases it may be necessary for the robot to produce explanations to make the domains compatible again (Chakraborti et al. 2017). Since model reconciliation is outside of the scope of legible planning, we will not cover the cases in which the two planning models require reconciliation, and to avoid reconciliation scenarios we introduce the hypothesis that every plan instance computed in $\mathbf{\Xi_{pl}}$ has exactly one valid description also in $\mathbf{\Xi'_{pl}}$.

Goal Recognition using PDDL

As previously discussed, we model the observer as a probabilistic goal recognizer $P(G|\Pi)P(\Pi)$ which infers the probability distribution over the possible goals given sequences of observed actions, possibly also integrating the observations with contextual information from its beliefs.

Though it is possible to implement it in many ways, we realized it by a probabilistic model based on the Planning Domain Description Language (PDDL). The provided formulation for goal recognition is based on (Persiani and Hellström 2020), and is flexible in how observed actions can appear in the plans $\Pi_{\Xi'}$, which, as we will later discuss, allows us to easily define a set of theory of mind models as functions which drop parts of the observations. In the following we give some technical implementation details.

PDDL (McDermott 1998) is a standard language to specify planning domains for what is usually referred to as *classical planning*. It is based on the STRIPS syntax and uses predicate logic to describe the current task state. In PDDL, a planning domain is specified by the tuple $\langle \mathcal{P}, \mathcal{A} \rangle$, where $\mathcal{P}$ is the set of possible truth predicates describing a state and $\mathcal{A}$ a set of operators that allow to transition between states. Every operator is defined by the triple $\langle pre(a), eff^-(a), eff^+(a) \rangle$. $pre(a)$ is a list of predicates that must be true in a given state for applying a to it, $eff^-(a)$ and $eff^+(a)$ are two lists of negative and positive effects which describe how the state is modified by a. For a specific planning domain $\langle \mathcal{P}, \mathcal{A} \rangle$, a derived planning instance is obtained by specifying the tuple $\Xi = \langle \mathcal{P}, \mathcal{A}, I, \mathcal{G} \rangle$. Where $I \subseteq \mathcal{P}$ is the initial state, $\mathcal{G} \subseteq \mathcal{P}$ is the target goal state. The goal of a planner is to find a valid sequence of operators $\pi \in \Pi$ that from I reaches $\mathcal{G}$ while incurring the least cost. From a planning domain $\langle \mathcal{P}, \mathcal{A} \rangle$ and a sequence of observations $\pi = \{o_0, ..., o_n\} \in \Pi$, goal recognition can be performed by providing a model for

$$P(G|\Pi)P(\Pi) = \beta P(\Pi|G)P(G) \tag{6}$$

Π is the set of valid partial plans inside Ξ, while $P(G)$ is the explicit prior probability of the goals. In this setting a possible way to realize $P(\Pi|G)$ is by computing, for every goal being considered, the cost of two optimal plans obtained from the planning instances $\Xi = \langle \mathcal{P}, \mathcal{A}, I, \mathcal{G} \rangle$ and $\Xi' = \langle \mathcal{P}, \mathcal{A}', I, \mathcal{G}' \rangle$, where:

- $\mathcal{A}' = \mathcal{A}$ with action effects modified as:
  $\forall a' \in \mathcal{A}'$
  - $pre(a') = pre(a) \wedge\ p_a$
  - $p_a = \wedge_i (x_{ai} = arg_{ai})$ if arg_{ai} is specified for action i
  - $\mathit{eff}^+(a') = \mathit{eff}^+(a) \wedge e_0$ if $a \in \pi$ and is the first of the observations (i.e. $n = 0$)
  - $\mathit{eff}^+(a') = \mathit{eff}^+(a) \wedge e_{n-1} \rightarrow e_n$ if $a \in \pi$ and $n \geq 1$
  - $\mathit{eff}^+(a') = \mathit{eff}^+(a)$ otherwise.
- $\mathcal{G}' = \mathcal{G} \cup e_n$, where e_n is the effect predicate of the last action in π.

The latter planning instance achieves $\mathcal{G}$ by producing a plan $\pi_{\Xi'}$ which is constrained to contain the observations π, the former instead achieves $\mathcal{G}$ by the means of an optimal plan π_Ξ. These two instances are used to evaluate the degree of rationality $R(\pi, \mathcal{G})$ that the observations have towards the possible goals $\mathcal{G}$, as computed by the formula:

$$R(\pi, \mathcal{G}) = \frac{|\pi_\Xi|}{|\pi_{\Xi'}|} \tag{7}$$

$|\pi'_\Xi|$ is the cost of the optimal plan achieving $\mathcal{G}$ using Ξ', while $|\pi_\Xi|$ is the cost of the optimal plan using Ξ. Given the hypothesis that an optimal plan π_Ξ exists, the following holds true:

$$|\pi_\Xi| \leq |\pi_{\Xi'}|, 0 \leq \frac{|\pi_\Xi|}{|\pi_{\Xi'}|} \leq 1 \tag{8}$$

All sequences of observations that induce plans with minimum cost have $R(\pi, \mathcal{G}) = 1$, while sub-optimal observation sequences towards $\mathcal{G}$ have $0 \leq R(\pi, \mathcal{G}) < 1$. Hence, the lower $R(\pi, \mathcal{G})$ is, the more sub-optimal it is to achieve that goal while being consistent with the observations. $R(\pi, \mathcal{G})$ captures a measure of rationality of the observed actions in the sense that it evaluates whether and how their investment of resources is efficient towards the available goals. By this definitions optimal plans are the also the most rationals (rationality is here equivalent as *predictability* in (Dragan and Srinivasa 2013) if we implicitly assume that the observer evaluates rational actions as predictable).

The probabilistic model for sequences of actions given the possible goals, $P(\Pi|G)$, can be finally obtained as a function of R, such as through a Boltzmann distribution:

$$\forall g \in G\ P(\pi|g) = \alpha e^{\frac{R(\pi,g)}{\tau}} \tag{9}$$

where α is the normalizing factor which marginalizes over all of the possible candidate goals, $\tau > 0$ is the distribution's temperature parameter. Given a sequence of observations, this model returns high probability for the goals that are rational to pursue, low probability otherwise.

Production of legible plans with off-the-shelf planners

As discussed in Section 3, legible planning means to find a plan π such that its goal is easily discernible from a set of other candidates in an observer perspective. This is obtained by making the observer model $P(G|\Pi)P(\Pi)$ to provide a

probability distribution over the goals as similar as possible to the one of the real goal with $P(G = \hat{g}) = 1$. With the introduction of the theory of mind T, Eq. 1 (and similarly Eq. 3,4) can be rewritten as:

$$\begin{aligned} \text{legibility}_T(\pi, g) &= \mathcal{H}(P_{g'}(G'), P(G'|\pi')P(\pi')) \\ g' = T(g), G' &= T(G), \pi' = T(\pi) \\ \{\pi, G\} \in \Xi_{pl}, &\{\pi', G'\} \in \Xi'_{pl}, \end{aligned} \tag{10}$$

where now the goal recognition is done utilizing the planning instance Ξ'_{pl} used by the observer to evaluate goals, which is obtained by applying T to Ξ_{pl}. Once inside Ξ'_{pl} goal recognition can be performed as previously described without further modifications. With a slight abuse of notation we use T to transform a planning instance in the broad sense and in particular also the plans and goals it yields.

Since plan legibility is non-monotonic (i.e. n–legibility$(\pi, g) \gtreqless (n+1)$–legibility$(\pi, g)$) in the general case, n-legibility requires to exhaustively search the plan space up to a depth of n to then compute the legibility at that level. In order to leverage already existing planners, we propose to utilize diverse planning techniques (Katz and Sohrabi 2020). Diverse planning is the task of finding k plans that achieve a target goal which are evaluated based on their cost and diversity. We use diverse planning as an alternative to search the plan space and use the parameter k to define its size. After that, the $\hat{n}$–legibility of a plan is found by iteratively computing its i–legibility values for $i = 1..n$, averaging the results to get the final value. Based on diverse planning and Eq. 10, our proposed algorithm to find the legible plan is shown in Algorithm 1.

Algorithm 1 $\hat{n}$-legible planning.

1: **procedure** LEGIBLE-PLANNING($\Xi, T, G, \hat{g}, \gamma, k$)
2: $\quad \Pi_{\hat{g}} \leftarrow$ DIVERSE-PLAN($\Xi, \hat{g}, k$)
3: $\quad \Xi' \leftarrow T(\Xi)$
4: $\quad \hat{g}' \leftarrow T(\hat{g})$
5: $\quad G' \leftarrow T(G)$
6: $\quad \overline{\pi} \leftarrow \underset{\pi \in \Pi_{\hat{g}}}{\text{argmax}}\ |\hat{T(\pi)}|$–legibility $- \gamma|\pi|$
7: **end procedure**

Algorithm 1 performs three main operations: first it computes the set of plans $\Pi_{\hat{g}}$ achieving the target goal in the planner's perspective, then it transforms the planning instance into its correspondings using the observer's model through the theory of mind T, finally, it finds the $|\hat{\pi}|$-legible plan also regularizing the result towards cheap plans, such that the efficiency of the returned plan can also be maintained. Legibility is computed in the observer's evaluated plan space $\Pi'_{\hat{g}}$, while the plan cost using the planner's task model.

Illustrative Example

To illustrate the production of legible plans we propose the following example in the Human-Robot teaming scenario shown in Figure 2. It is a rescue scenario where an autonomous scout robot is moving in a dangerous environment (e.g. a laboratory filled with toxic gases) searching for people, while a commander supervisions the operations on its computer interface from outside of the building. Both have access to a map of the environment in the form of a grid-world on which the robot's position is being tracked. We assume the commander doesn't have direct access to the internal goal of the autonomous robot, nor of other map features such as obstacles or rough terrain, which instead are detectable by the robot through its sensors. When the mission starts the robot communicate its goal explicitly to the commander through e.g. a speech interface, but in order to continuously and implicitly better communicate its goal, the autonomous robot estimates the commander's other expected possible goals, and produce a legible plan such that its true goal is best discriminated.

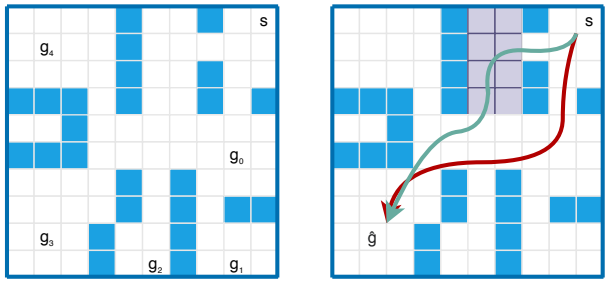

Figure 2: Legible planning in the rescue scenario. Left side: world grid that the commander believes the robot is planning for. This plan instance is estimated internally by the robot through a theory of mind, and doesn't have any obstacles. Right side: real world grid computed by the robot, which has the real goal and contains detected obstacles.

To make its plan more legible from the commander perspective, the robot should evaluate the possible candidate goals that its actions might communicate to the commander as being pursued (left side of Figure 2). In this case a possible estimate is that the commander thinks that all of the rooms of the building are candidate destinations for the plans. In this setting, the right side of Figure 2 compares the legible plan (green) with the optimal plan (red) for the same true goal g_3, computed in the robot perspective. The legible plan is more expensive, as it passes over some obstacles (purple), yet since it avoids as much as possible going towards other candidates goals, it results being more legible by the commander. It better communicates that g_3 is the true pursued goal, and symmetrically, that the other candidate goals are not the ones being pursued.

The commander's perspective is estimated by the scout robot through its theory of mind. In this particular scenario the robot estimates that the commander believes the robot can move freely and that the environment is unobstructed, which is plausible in this scenario as only the robot can map the real environment. Since everything is computed locally to the robot, the commander is not required to perform any operation to communicate its mental state in the form of an

Domain	$\lvert\mathcal{A}\rvert$	$\lvert\mathcal{P}\rvert$	$\lvert I\rvert$	$\lvert G\rvert$	$\lvert\overline{\pi}\rvert$
intrusion	9.0	11.0	1.0	4.75	17.15
kitchen	29.0	23.0	2.0	1.0	10.6
satellite	5.0	12.0	62.8	6.8	16.55
campus	22.0	12.0	1.0	2.75	4.925
blocks-world	4.0	5.0	14.4	4.95	15.25
logistics	6.0	3.0	22.7	2.3	31.25
easy-ipc-grid	3.0	8.0	227.4	1.0	17.2
miconic	4.0	8.0	518.6	6.6	24.85
ferry	3.0	7.0	99.3	8.9	28.27

Table 1: Average instance measures over the tested planning domains. The columns, from left to right are: number of operators, number of predicates, size of the initial state, size of the goal, length of optimal plan.

expected planning instance, such as through annotating the map.

Evaluation

In order to benchmark our method on PDDL domains, we run Algorithm 1 on the following planning domains: *logistics, blocks-world, intrusion-detection, kitchen, campus, satellite, easy-ipc-grid, miconic, ferry* (Pereira, Oren, and Meneguzzi 2017)[1], selecting for every domain 10 random planning instances. Table 1 shows average relevant metrics of the instances. To compute plans and perform goal recognition we utilize the *Forbid-Iterative*[2] (Katz and Sohrabi 2020) planner which is based on *Fast-Downward* (Helmert 2006) and performs diverse planning by iteratively creating plan instances forbidding previously found plans. The distance of the true goal distribution and the observer's predicted distribution is realized through the cross-entropy function.

For every planning instance we applied Algorithm 1 by the following procedure: we selected one goal randomly chosen amongst a set of 4 candidates as true goal for that instance, and used diverse planning to compute $k = 50, 100, 200$ plans toward it. For every plan its resulting legibility value is found by averaging the legibility values for all of its steps.

In the general case, planner and observer have two disjoint task models which makes their theory of mind non trivial and for which a qualitative evaluation would require an initial learning phase and model reconciliation to align them. However, since we want to test many PDDL domains, we want to evaluate many possible theory of minds statistically rather than qualitatively. In this setting we utilize a subset of all of the possible theory of minds that is easy to randomly sample. We focus more in particular on the theory of minds which treat the transformation between planning instances T as a function dropping parts of the actions or parameters. This corresponds in estimating the observer as having strictly less or equal information as the planner when inferring goals.

More precisely we don't assume, in the inferred plans, for the observations to appear in the exact succession in which they were gathered, but other actions are allowed to appear in between them. All of the observations must nevertheless appear in the observed order in their corresponding inferred plans. This simulates the observer not being able to recognize all of the actions. Additionally, we allow for observations to be partially instantiated, i.e. not all of their parameters are required to be specified, with missing parameters being rather inferred during plan recognition. This simulates how the observer doesn't have access to all of the information contained in the observed actions and infers the missing parts. Even though we find it plausible also in real implementations, we selected this family of theory of mind because it is easy to simulate and sample, which each sampled T randomly dropping $d\%$ of the plans actions and parameters. The tested values of $d\%$ are $0, 20, 40, 60, 80$.

For every planning instance the main gathered measures are the *legibility gain* and the *cost gain*. The legibility gain $\mathcal{L}_{gain}$ corresponds to the ratio between the legibility of the optimal plan $\overline{\pi}$ and the one of the most legible plan among the k generated plans, π_{legible}.

$$\mathcal{L}_{gain} = \frac{\hat{n}\text{-legibility}(\pi_{\text{legible}})}{\hat{n}\text{-legibility}(\overline{\pi})} \tag{11}$$

The legibility gain provides indication on how advantageous it is, in terms of legibility, to follow a legible plan rather than the optimal plan. The total legibility of a plan is computed by averaging all of the legibilities of its steps with equal weight. The other relevant indicator is the cost gain $\mathcal{C}_{gain}$, which instead indicates how costly (in plan length) the legible plan is when compared with the optimal plan.

$$\mathcal{C}_{gain} = \frac{|\pi_{\text{legible}}|}{|\overline{\pi}|} \tag{12}$$

These two measures together show the trade-off between plan legibility and cost. Table 2 illustrates the measures gathered on the tested domains. The obtained values for a domain are the averages of all the 10 instances belonging to it.

Discussion

Our measurements show a positive correlation between legibility and cost: it was always possible to increase legibility in exchange of making plans more lengthy (see Table 3). The measured gains seems however to be strongly dependent on the tested domain. This is better highlighted in Figure 3, where legibility is compared with the percentage of random items being dropped from the plans. For example, in *logistics* the legibility peaks when none of the observations are dropped, while for *intrusion-detection* legibility increases linearly in the opposite direction. Dropping parts of observed plans relaxes plan recognition in the sense that more possible plans fits the same observations, which translates in making the sequences of observations more probable

[1] *https://github.com/pucrs-automated-planning/goal-plan-recognition-dataset*. Accessed August 12, 2021.

[2] *https://github.com/IBM/forbiditerative*. Accessed August 12, 2021.

		γ							
		0		0.001		0.01		0.1	
$d\%$	*domain*	$\mathcal{L}_{gain}$	$\mathcal{C}_{gain}$	$\mathcal{L}_{gain}$	$\mathcal{C}_{gain}$	$\mathcal{L}_{gain}$	$\mathcal{C}_{gain}$	$\mathcal{L}_{gain}$	$\mathcal{C}_{gain}$
0	*campus*	1.32	1.1	1.31	1.1	1.26	1.1	1.08	1.03
	kitchen	1.43	1.26	1.37	1.26	1.16	1.11	1.0	1.0
	satellite	1.38	1.34	1.35	1.32	1.2	1.2	1.01	1.01
	intrusion	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
	ferry	1.58	1.3	1.48	1.29	1.16	1.08	1.02	1.02
	logistics	3.08	3.21	2.33	2.39	1.43	1.17	1.06	1.01
	blocks	2.44	6.24	1.87	6.24	1.24	2.54	1.0	1.0
	grid	1.03	1.35	1.02	1.29	1.0	1.22	1.0	1.0
	miconic	1.08	1.2	1.07	1.13	1.04	1.05	1.01	1.0
20	*campus*	1.25	1.05	1.25	1.05	1.21	1.05	1.08	1.02
	kitchen	1.58	1.13	1.52	1.11	1.29	1.06	1.02	1.05
	satellite	1.28	1.29	1.26	1.26	1.15	1.19	1.01	1.01
	intrusion	1.07	1.04	1.07	1.04	1.04	1.04	1.0	1.0
	ferry	1.99	1.25	1.89	1.25	1.45	1.15	1.04	1.05
	logistics	1.7	1.44	1.64	1.42	1.33	1.25	1.04	1.01
	blocks	3.03	5.84	2.24	5.04	1.33	2.39	1.02	1.0
	grid	1.05	1.41	1.04	1.38	1.02	1.27	1.0	1.0
	miconic	1.13	1.16	1.12	1.16	1.07	1.03	1.02	1.0
40	*campus*	1.19	1.15	1.18	1.15	1.15	1.08	1.04	1.02
	kitchen	1.51	1.11	1.47	1.08	1.28	1.06	1.02	1.05
	satellite	1.19	1.3	1.17	1.24	1.1	1.13	1.01	1.0
	intrusion	1.12	1.08	1.12	1.08	1.07	1.06	1.0	1.01
	ferry	1.44	1.45	1.4	1.42	1.2	1.17	1.02	1.0
	logistics	1.53	1.54	1.47	1.49	1.24	1.27	1.04	1.01
	blocks	2.91	5.75	2.2	5.13	1.3	2.02	1.03	1.0
	grid	1.09	1.5	1.08	1.5	1.04	1.3	1.0	1.0
	miconic	1.12	1.19	1.11	1.19	1.07	1.03	1.02	1.0
60	*campus*	1.2	1.15	1.2	1.15	1.16	1.13	1.06	1.01
	kitchen	1.63	1.16	1.59	1.16	1.36	1.11	1.05	1.05
	satellite	1.14	1.25	1.13	1.19	1.08	1.13	1.01	1.0
	intrusion	1.17	1.35	1.15	1.28	1.08	1.07	1.0	1.01
	ferry	1.23	1.41	1.2	1.41	1.1	1.13	1.01	1.0
	logistics	1.31	1.59	1.28	1.44	1.14	1.17	1.01	1.0
	blocks	2.14	5.76	1.76	4.95	1.17	1.82	1.02	1.0
	grid	1.11	1.87	1.1	1.82	1.04	1.49	1.0	1.0
	miconic	1.06	1.19	1.05	1.16	1.03	1.02	1.01	1.0
80	*campus*	1.2	1.17	1.2	1.17	1.16	1.17	1.05	1.01
	kitchen	1.47	1.15	1.44	1.15	1.3	1.11	1.05	1.05
	satellite	1.1	1.27	1.09	1.27	1.05	1.08	1.01	1.0
	intrusion	1.16	1.37	1.14	1.35	1.07	1.11	1.0	1.01
	ferry	1.09	1.37	1.08	1.23	1.02	1.09	1.0	1.0
	logistics	1.15	1.44	1.13	1.32	1.06	1.1	1.01	1.0
	blocks	1.59	5.83	1.43	4.49	1.11	1.65	1.02	1.0
	grid	1.12	2.59	1.1	2.31	1.04	1.47	1.0	1.0
	miconic	1.06	1.15	1.05	1.14	1.03	1.0	1.01	1.0

Table 2: Computed plan legibility and cost measures. $\mathcal{L}_{gain}$: cost gain, $\mathcal{C}_{gain}$: legibility gain, $d\%$: percentage of dropped observations, γ: regularization factor towards cheap plans. Values are for $k = 200$.

toward all goals simultaneously. Our hypothesis is that planning domains with goal-specific actions will benefit more from dropping part of the observations rather than domains with universally applicable actions, as in the former case relaxed plans are still bound by the goal-specific actions. This currently gathered data indicates that it is difficult to predict a priori how legibility will behave in experiments on real data.

The results further show that the regularization factor γ plays an important role. In domains such as *blocks-world* or *logistics* some action loops can arbitrarily increase the legibility towards the goals. This happens to the expense of a huge drop of performance of the plans. In those domains, without regularization legible plans could reach a length up to 3-6 times higher than the optimal plans for the same instances.

Conclusions

In this paper we proposed a procedure to compute legible plans using off-the-shelf PDDL planners. Our formulation is based on a probabilistic formulation of goal recognition

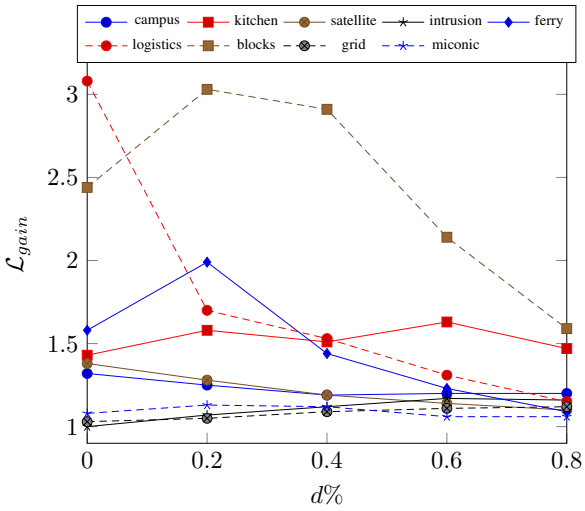

Figure 3: Legibility gain for various percentages of dropped parameters.

	k	50	100	200
$\gamma = 0$	$\mathcal{L}_{gain}$	1.31	1.43	1.54
	$\mathcal{C}_{gain}$	1.36	1.47	2.09
$\gamma = 0.01$	$\mathcal{L}_{gain}$	1.13	1.18	1.16
	$\mathcal{C}_{gain}$	1.18	1.21	1.36
$\gamma = 0.1$	$\mathcal{L}_{gain}$	1.01	1.02	1.01
	$\mathcal{C}_{gain}$	1	1	1

Table 3: Average over all of the domains of legibility and cost gains for increasing values of k.

and models the observer's task model as being estimated by the planner. The integration of these two task models is obtained by a procedure based on theory of mind which transforms the planner's task model into the same from the observer's perspective. The introduction of theory of mind for legibility is crucial as it describes how planner and observer tasks models are connected, and makes the planner to know how its actions are perceived.

We further proposed an illustrative example based on a rescue scenario with a robot and a human commander explaining how legibility behaves in this context. In Human-Robot Teaming contexts we would like to highlight that since legibility leverages the theory of mind owned by the robot, it doesn't require inputs by the commander such as labeling of the scene. This is particularly interesting in the context of highly autonomous robots. Additionally, a statistical evaluation over several planning domains showed how the proposed algorithm successfully generates plans that are more legible than the optimal plans, also indicating how legibility is a trade-off with plan cost. Our investigations show that this relation between legibility and cost is strongly dependent on the domain and theory of mind being utilized. In our statistical experiments we tested the family of theory of minds that strictly reduce the amount of information for goal recognition.

Apart from illustrating the broad possibility of increasing legibility of plans, these current tests cannot reach further conclusions a priori valid on every possible domain, which rather require investigation on the specific domain and theory of mind being utilized.

Future work regards the investigation of legibility as perceived by humans in real scenarios. In particular on whether the proposed model for plan legibility, which is based on the concept of rationality, effectively capture human judgement of legibility in domains specific for human-robot interaction.

References

Chakraborti, T.; Sreedharan, S.; Zhang, Y.; and Kambhampati, S. 2017. Plan explanations as model reconciliation: Moving beyond explanation as soliloquy. *arXiv preprint arXiv:1701.08317*.

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 international conference on automated planning and scheduling*, volume 29, 86–96.

Devin, S., and Alami, R. 2016. An implemented theory of mind to improve human-robot shared plans execution. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 319–326. IEEE.

Dragan, A., and Srinivasa, S. 2013. Generating legible motion.

Dragan, A., and Srinivasa, S. 2014. Integrating human observer inferences into robot motion planning. *Autonomous Robots* 37(4):351–368.

Dragan, A. D.; Bauman, S.; Forlizzi, J.; and Srinivasa, S. S. 2015. Effects of robot motion on human-robot collaboration. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 51–58. IEEE.

Dragan, A. D.; Lee, K. C.; and Srinivasa, S. S. 2013. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 301–308. IEEE.

Fox, M.; Long, D.; and Magazzeni, D. 2017. Explainable planning. *arXiv preprint arXiv:1709.10256*.

Hellström, T., and Bensch, S. 2018. Understandable robots - What, Why, and How. *Paladyn, Journal of Behavioral Robotics* 9(1):110–123.

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

Katz, M., and Sohrabi, S. 2020. Reshaping diverse planning. In *AAAI*, 9892–9899.

Kulkarni, A.; Zha, Y.; Chakraborti, T.; Vadlamudi, S. G.; Zhang, Y.; and Kambhampati, S. 2019. Explicable planning as minimizing distance from expected behavior. In *AAMAS*, 2075–2077.

Kulkarni, A.; Srivastava, S.; and Kambhampati, S. 2019. A unified framework for planning in adversarial and cooperative environments. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 2479–2487.

Kulkarni, A.; Srivastava, S.; and Kambhampati, S. 2020. Signaling friends and head-faking enemies simultaneously: Balancing goal obfuscation and goal legibility. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, 1889–1891.

MacNally, A. M.; Lipovetzky, N.; Ramirez, M.; and Pearce, A. R. 2018. Action selection for transparent planning. In *AAMAS*, 1327–1335.

McDermott, D. 1998. Pddl-the planning domain definition language.

Meijering, B.; Van Rijn, H.; Taatgen, N.; and Verbrugge, R. 2011. I do know what you think i think: Second-order theory of mind in strategic games is not that difficult. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 33.

Miura, S., and Zilberstein, S. 2020. Maximizing plan legibility in stochastic environments. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, 1931–1933.

Pereira, R. F.; Oren, N.; and Meneguzzi, F. 2017. Landmark-based heuristics for goal recognition. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Persiani, M., and Hellström, T. 2020. Intent recognition from speech and plan recognition. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, 212–223. Springer.

Zhang, Y.; Sreedharan, S.; Kulkarni, A.; Chakraborti, T.; Zhuo, H. H.; and Kambhampati, S. 2017. Plan explicability and predictability for robot task planning. In *2017 IEEE international conference on robotics and automation (ICRA)*, 1313–1320. IEEE.