Explicable Planning as Minimizing Distance from Expected Behavior[*][†]

Extended Abstract

Anagha Kulkarni
Arizona State University
anaghak@asu.edu

Yantian Zha
Arizona State University
yzha3@asu.edu

Tathagata Chakraborti
IBM Research AI
tchakra2@ibm.com

Satya Gautam Vadlamudi
CreditVidya
gautam.vadlamudi@creditvidya.com

Yu Zhang
Arizona State University
yzhan442@asu.edu

Subbarao Kambhampati
Arizona State University
rao@asu.edu

ABSTRACT

In order to achieve effective human-AI collaboration, it is necessary for an AI agent to align its behavior with the human's expectations. When the agent generates a task plan without such considerations, it may often result in *inexplicable behavior* from the human's point of view. This may have serious implications for the human, from increased cognitive load to more serious concerns of safety around the physical agent. In this work, we present an approach to generate explicable behavior by minimizing the distance between the agent's plan and the plan expected by the human. To this end, we learn a mapping between plan distances (distances between expected and agent plans) and human's plan scoring scheme. The plan generation process uses this learned model as a heuristic. We demonstrate the effectiveness of our approach in a delivery robot domain.

ACM Reference Format:
Anagha Kulkarni, Yantian Zha, Tathagata Chakraborti, Satya Gautam Vadlamudi, Yu Zhang, and Subbarao Kambhampati. 2019. Explicable Planning as Minimizing Distance from Expected Behavior. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 3 pages.

1 EXPLICABLE PLANNING

A robot's behavior should not only be optimal with respect to its own model, but also explicable with respect to a human's mental model of the robot. The problem of inexplicable behavior arises when the robot's plan deviates from that expected by the human. In this paper, we explore the plan explicability problem in a setting where the robot has access to a generative model of the human expectations. Even with a known mental model this remains a challenging and nuanced problem since the mental model might entail plans that are infeasible or prohibitively expensive for the robot, and thus at best can serve as a guide, and not an oracle, for generating explicable plans. The problem of generating explicable plans can be also solved in a model-free setting, where the robot does not have access to the human's mental model [7].

A Classical Planning Problem [3, 5] is a tuple $\mathcal{P} = \langle \mathcal{M}, \mathcal{I}, \mathcal{G} \rangle$, where $\mathcal{M} = \langle \mathcal{F}, \mathcal{A} \rangle$ is the domain model, that consists of a finite set of fluents, $\mathcal{F}$, and a set of actions, $\mathcal{A}$. A state s of the world is an instantiation of all fluents in $\mathcal{F}$. Let $\mathcal{S}$ be the set of states. $\mathcal{I} \subseteq \mathcal{S}$ is the initial state. $\mathcal{G}$ is the goal where a subset of fluents in $\mathcal{F}$ are instantiated. Each action $a \in \mathcal{A}$ is a tuple $\langle pre_a, add_a, del_a, c_a \rangle$ where c_a is the cost of a, $pre_a, add_a, del_a \subseteq \mathcal{F}$ are the preconditions, add and delete effects of a. $\Gamma(\cdot)$ is the transition function, such that, $\Gamma(s, a) \models \perp$, if $s \not\models pre_a$; else $\Gamma(s, a) \models s \cup add_a \setminus del_a$. The solution to the planning problem is a *plan* or a sequence of actions $\pi = \langle a_1, a_2, \ldots, a_n \rangle$ such that starting from the initial state, by sequentially executing the actions the agent achieves its goal, i.e. $\Gamma(\mathcal{I}, \pi) \models \mathcal{G}$. The cost of the plan is given by, $c(\pi) = \sum_{a_i \in \pi} c(a_i)$. An optimal plan achieves the goal with minimum cost.

An Explicable Planning Problem is defined as a tuple $\mathcal{P}_{EPP} = \langle \mathcal{M}^{\mathcal{R}}, \mathcal{M}^{\mathcal{R}}_{\mathcal{H}}, \mathcal{I}^{\mathcal{R}}, \mathcal{G}^{\mathcal{R}} \rangle$, where, $\mathcal{M}^{\mathcal{R}} = \langle \mathcal{F}^{\mathcal{R}}, \mathcal{A}^{\mathcal{R}} \rangle$ is the robot's domain model, such that, $\mathcal{F}^{\mathcal{R}}$ and $\mathcal{A}^{\mathcal{R}}$ represent the robot's fluents and actions, $\mathcal{M}^{\mathcal{R}}_{\mathcal{H}} = \langle \mathcal{F}^{\mathcal{R}}, \mathcal{A}^{\mathcal{R}}_{\mathcal{H}} \rangle$ is the human's mental model of the robot where $\mathcal{F}^{\mathcal{R}}$ and $\mathcal{A}^{\mathcal{R}}_{\mathcal{H}}$ represent the fluents and actions that the human thinks are available to the robot, $\mathcal{I}^{\mathcal{R}}$ and $\mathcal{G}^{\mathcal{R}}$ are the initial and goal states of the robot. $\mathcal{A}^{\mathcal{R}}$ and $\mathcal{A}^{\mathcal{R}}_{\mathcal{H}}$ represent that the action names, preconditions, effects and costs of the actions can be different. The initial state and the goal state are assumed to be known to the human. Further, we define an evaluation function δ^* that represents the difference between the robot plan and the expected plan and thus the cost of being inexplicable.

An Explicable Plan $\pi^*_{\mathcal{M}^{\mathcal{R}}}$ is a solution to $\mathcal{P}_{EPP}$ that minimizes the sum of the plan cost and the evaluation function:

$$\pi^*_{\mathcal{M}^{\mathcal{R}}} = \arg\min_{\pi_{\mathcal{M}^{\mathcal{R}}}} c(\pi_{\mathcal{M}^{\mathcal{R}}}) + \delta^*(\pi_{\mathcal{M}^{\mathcal{R}}}, \pi_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}})$$

2 GENERATION OF EXPLICABLE PLANS

Generation of explicable plans involves the following 4 steps: **(1)** Since the evaluation function is not directly available to the robot, we learn an approximation of it using a combination of three plan distance measures – action, causal link and state sequence distances – proposed by Nguyen et al. [4], Srivastava et al. [6]. **(2)** Human evaluators are asked to score each action in candidate robot plans with 1 (explicable) or 0 (inexplicable). The total explicability score of each plan is computed as an average of explicable actions. **(3)** A mapping between the three plan distances and the explicability

[*] The full version of the paper is available at https://arxiv.org/abs/1611.05497.
[†] A detailed treatise of explicable planning and its relation to other forms of explainable robot behavior can be found in [1].

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

Algorithm 1 Reconciliation Search

```
Input: $\mathcal{P}_{\mathcal{EPP}} = \langle \mathcal{M}^{\mathcal{R}}, \mathcal{M}^{\mathcal{R}}_{\mathcal{H}}, \mathcal{I}^{\mathcal{R}}, \mathcal{G}^{\mathcal{R}} \rangle$, $max\_cost$, $Exp\,(\cdot)$
Output: $\mathcal{E}_{\mathcal{EPP}}$ (set of explicable plans)
$\mathcal{E}_{\mathcal{EPP}} \leftarrow \emptyset$; $open \leftarrow \emptyset$; $closed \leftarrow \emptyset$
$open$.insert($\mathcal{I}$, 0, inf)
while $open \neq \emptyset$ do
    $n \leftarrow open$.remove()                    ▷ Node with highest $h(\cdot)$
    if $n \models \mathcal{G}$ then
        $\mathcal{E}_{\mathcal{EPP}}$.insert($\pi$ s.t. $\Gamma_{\mathcal{M}^{\mathcal{R}}}(\mathcal{I}, \pi) \models n$)
    $closed$.insert($n$)
    for each $v \in$ successors($n$) do
        if $v \notin closed$ then
            if $g(n) + cost(n, v) \leq max\_cost$ then
                $open$.insert($v$, $g(n) + cost(n, v)$, $h(v)$)
        else
            if $h(n) < h(v)$ then
                $closed$.remove($v$)
                $open$.insert($v$, $g(n) + cost(n, v)$, $h(v)$)
return $\mathcal{E}_{\mathcal{EPP}}$
```

score for each candidate plan, referred to as the *explicability distance*, is learned using a regression model. **(4)** In order to generate explicable plans, the plan score predictions provided by the explicability distance are used as a heuristic to guide an anytime search implemented in `Fast-Downward` [2] planner. The search produces increasingly explicable plans in an anytime fashion.

The Explicability Distance approximates the evaluation function δ^* using a combination of three plan distances. It is a regression model that is trained to learn the scoring scheme of the users by mapping the plan scores to the corresponding *explicability feature vector* as described below.

The explicability feature vector for a candidate robot plan is constructed by computing the plan's distance from an expected plan. We define a set of expected plans, $\mathcal{E}(\mathcal{P})$, for the planning problem $\mathcal{P}$ as the set of optimal solutions to it. $\mathcal{E}(\mathcal{P}^{\mathcal{R}}_{\mathcal{H}})$ denotes the set of optimal plans to $\mathcal{P}^{\mathcal{R}}_{\mathcal{H}}$ that the human expects the robot to compute. An expected plan in $\mathcal{E}(\mathcal{P}^{\mathcal{R}}_{\mathcal{H}})$ that is closest to the candidate robot plan is used to construct the feature vector. It is represented as $\pi^*_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}$, and is found as follows:

$$\pi^*_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}} = \{\pi_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}} \mid \underset{\pi_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}} \in \mathcal{E}(\mathcal{P}^{\mathcal{R}}_{\mathcal{H}})}{\arg\min} \quad \delta_{exp}(\pi_{\mathcal{M}^{\mathcal{R}}}, \pi_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}})\}$$

where δ_{exp} is the distance between the robot plan $\pi_{\mathcal{M}^{\mathcal{R}}}$ and an expected plan $\pi_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}$. δ_{exp} is computed using the three plan distance measures as follows:

$$\begin{aligned}\delta_{exp}(\pi_{\mathcal{M}^{\mathcal{R}}}, \pi_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}) = ||&\delta_A(\pi_{\mathcal{M}^{\mathcal{R}}}, \pi_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}) \\ &+ \delta_C(\pi_{\mathcal{M}^{\mathcal{R}}}, \pi_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}) + \delta_S(\pi_{\mathcal{M}^{\mathcal{R}}}, \pi_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}})||_2\end{aligned}$$

where δ_A, δ_C, δ_S represent action, causal link and state sequence distances respectively. The explicability feature vector, Δ, is:

$$\Delta = \langle \delta_A(\pi_{\mathcal{M}^{\mathcal{R}}}, \pi^*_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}), \delta_C(\pi_{\mathcal{M}^{\mathcal{R}}}, \pi^*_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}), \delta_S(\pi_{\mathcal{M}^{\mathcal{R}}}, \pi^*_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}) \rangle$$

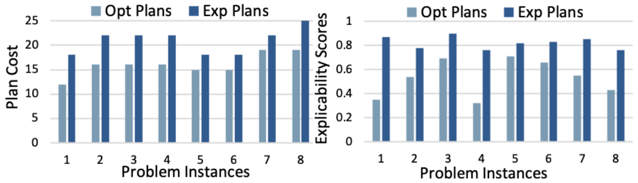

Figure 1: Comparison of (a) plans costs and (b) explicability scores provided by subjects for optimal and explicable plans.

Finally, the explicability distance $Exp\,(\pi_{\mathcal{M}^{\mathcal{R}}} \;/\; \pi^*_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}})$ is a regression function f that fits the three plan distances to the explicability scores of the robot plans:

$$Exp\,(\pi_{\mathcal{M}^{\mathcal{R}}} \;/\; \pi^*_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}) \approx f(\Delta, b), \text{ where } b \text{ is the parameter vector.}$$

The plan generation process outlined in Algorithm 1 employs a cost-bounded anytime greedy search that generates all valid loopless candidate plans up to a given cost bound and then progressively searches for plans with better explicability scores. We use $Exp\,(\cdot)$ as a heuristic to guide our search. The heuristic value $h(v)$ of a state v is computed using the agent plan prefix $\pi_{\mathcal{M}^{\mathcal{R}}}$ leading up to v:

$$\begin{aligned}&h(v) \;=\; Exp\,(\pi_{\mathcal{M}^{\mathcal{R}}} / \pi'_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}) \\ &\text{where } \Gamma_{\mathcal{M}^{\mathcal{R}}}(\mathcal{I}, \pi_{\mathcal{M}^{\mathcal{R}}}) \models v \;\wedge\; \Gamma_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}(\mathcal{I}, \pi'_{\mathcal{M}^{\mathcal{R}}_{\mathcal{H}}}) \models v\end{aligned}$$

Property: *Explicablity score of a plan is non-monotonic.* This is because, as a partial plan grows, a new action may increase or decrease the score. $Exp\,(\cdot)$ in turn also exhibits non-monotonicity. The proposed anytime search comes in handy particularly in light of this non-monotonicity by being able to produce solutions with increasing degrees of explicability.

Demonstration: We designed a delivery robot domain to demonstrate explicable behaviors using a robot. The robot can deliver parcels/electronic devices and serve beverages to the humans using a tray. Whenever the robot carries the beverage cup there is some risk that the cup may tip over and spill the contents over the electronic items on the tray. Here the robot has to learn the context of carrying devices and beverages separately even if it results in an expensive plan (in terms of cost or time). A video of the demonstration can be viewed at https://bit.ly/2JweeYk.

Evaluation: Figure 1a shows that, for this domain, all explicable plans are more expensive than optimal plans. This is because the explicable plans involve doing separate trips for items of different types. A cost-optimal planner would not generate these plans. Figure 1b shows the results of a user study where the test subjects had to score the explicability of the plans. The explicability scores provided by the subjects are higher for explicable plans. *Further details are available in the full report.*[*]

ACKNOWLEDGMENTS

This research is supported in part by the ONR grants N00014-16-1-2892, N00014-18-1-2442, N00014-18-1-2840, the AFOSR grant FA9550-18-1-0067, the NASA grant NNX17AD06G and the NSF grant IIS-1844524.

REFERENCES

[1] Tathagata Chakraborti, Anagha Kulkarni, Sarath Sreedharan, David E. Smith, and Subbarao Kambhampati. 2019. Explicability? Legibility? Predictability? Transparency? Privacy? Security? The Emerging Landscape of Interpretable Agent Behavior. *ICAPS* (2019).

[2] Malte Helmert. 2006. The Fast Downward Planning System. *Journal of Artificial Intelligence Research (JAIR)* (2006).

[3] Drew McDermott, Malik Ghallab, Adele Howe, Craig Knoblock, Ashwin Ram, Manuela Veloso, Daniel Weld, and David Wilkins. 1998. PDDL – The Planning Domain Definition Language. (1998).

[4] Tuan Anh Nguyen, Minh Do, Alfonso Emilio Gerevini, Ivan Serina, Biplav Srivastava, and Subbarao Kambhampati. 2012. Generating Diverse Plans to Handle Unknown and Partially Known User Preferences. *Artificial Intelligence* (2012).

[5] Stuart J. Russell and Peter Norvig. 2003. *Artificial Intelligence: A Modern Approach* (2 ed.). Pearson Education.

[6] Biplav Srivastava, Tuan Anh Nguyen, Alfonso Gerevini, Subbarao Kambhampati, Minh Binh Do, and Ivan Serina. 2007. Domain Independent Approaches for Finding Diverse Plans. In *IJCAI*.

[7] Yu Zhang, Sarath Sreedharan, Anagha Kulkarni, Tathagata Chakraborti, Hankz H Zhuo, and Subbarao Kambhampati. 2017. Plan Explicability and Predictability for Robot Task Planning. In *ICRA*.