Plan Explicability for Robot Task Planning

Yu Zhang, Sarath Sreedharan, Anagha Kulkarni, Tathagata Chakraborti,
Hankz Hankui Zhuo and Subbarao Kambhampati
Email: {yzhan442, ssreedh3, akulka16, tchakra2, rao}@asu.edu, zhuohank@mail.sysu.edu.cn

***Abstract*—A desirable capability of agents is to respond to goal-oriented commands by autonomously constructing task plans. However, such autonomy can add significant cognitive load and potentially introduce safety risks to humans when agents behave unexpectedly. Hence, one important requirement is for such agents to synthesize plans that can be easily understood by humans. While there exists previous work that studied socially acceptable robots that interact with humans in "natural ways", and work that investigated legible motion planning, there exists no general solutions for high level task planning. To address this issue, we introduce the notion of plan *explicability*. To compute this measure for plans, first, we postulate that humans understand agent plans by associating abstract tasks with agent actions, which can be considered as a labeling process. We learn the labeling scheme of humans for agent plans from training examples using conditional random fields (CRFs). Then, we use the learned model to label a new plan to compute its explicability, which are used to guide the search.**

I. Introduction

Significant research efforts have been invested to build robotic agents that are more autonomous. These agents respond to goal specifications instead of basic motor commands, which requires them to autonomously synthesize task plans and execute those plans to achieve the goals. However, if the behaviors of these agents are incomprehensible, it can increase the cognitive load of humans and potentially introduce safety risks to them. As a result, one important requirement for such intelligent agents is to ensure that the synthesized plans are comprehensible to humans. This means that instead of considering only the planning model of the agent, plan synthesis should also consider the interpretation of the agent behavior from the human's perspective. This interpretation is related to our modeling of other agents. More specifically, we tend to have expectations of others' behaviors based on our understanding (modeling) of their capabilities, mental states and etc. If their behaviors do not match with these expectations, we would often be confused. **The reason for this confusion is due to the fact that our understanding of others' models is often different from the actual models**. This is particularly true when humans interact with robots since it is likely that we only have a partial (and potentially inaccurate) model of these intelligent agents.

For example, to darken a room that is too bright, a robot can either adjust the window blinds, switch off the lights, or break the light bulbs in the room. While breaking the light bulbs may well be the least costly plan to the robot under certain conditions (e.g., when the robot cannot easily move in the environment but we are unaware of it), it is clear that the other two options are far more desirable in the context of robots cohabiting with humans.

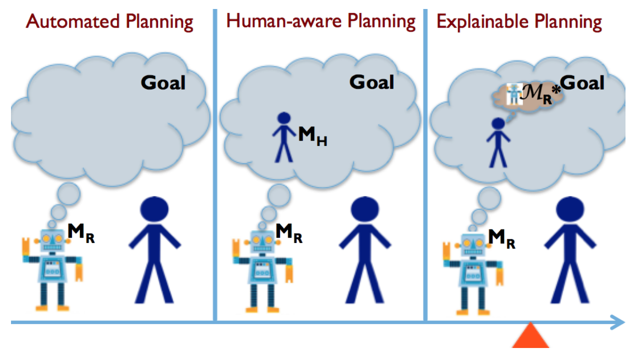

Fig. 1. From left to right, the scenarios illustrate the differences between automated task planning, human-aware planning and explicable planning (this work). In human-aware planning, the robot needs to maintain a model of the human (i.e., M_H) which captures the human's capabilities, intents and etc. In explicable planning, the robot considers the differences between its model from the human's perspective (i.e., $\mathcal{M}_R^*$) and its own model M_R.

In this paper, we introduce the notion of plan *explicability*, which is used by autonomous agents (e.g., robots) to synthesize "explicable plans" that can be easily understood by humans. As suggested in psychological studies [12, 5], we assume that humans naturally interpret a plan as achieving abstract tasks (or subgoals), which are functional interpretations of agent action sequences in the plan. For example, a robot that executes a sequence of manipulation actions may be interpreted as achieving the task of "$picking\ up\ cup$". Based on this assumption, intuitively, the easier it is for humans to associate tasks with actions in a plan, the more explicable the plan is. Since the association between tasks and agent actions can be considered as a labeling process, we learn the labeling scheme of humans for agent plans from training examples using conditional random fields (CRFs). We then use the learned model to label a new plan to compute its explicability, which is used by agents to synthesize plans that are more explicable without affecting the quality much.

II. Related Work

A planning capability allows agents to autonomously synthesize plans to achieve a goal given the agent model (M_R in Fig. 1). However, to work alongside of humans, these agents must be "human-aware" when synthesizing plans. In prior research, this issue is addressed under human-aware planning [4, 3, 11, 13] in which agents take the human's activities and intents into account when constructing their plans. This corresponds to human modeling in human-aware planning as shown in the second scenario in Fig. 1.

While our work on plan explicability falls within the scope of human-in-the-loop planning (which includes human-aware planning), it differs significantly from the previous work. In human-aware planning, the challenge is to obtain the human model (M_H in Fig. 1) which captures human capabilities [14], intents [11, 3] and etc. The modeling here is one level deeper: it is about the interpretation of the agent model from the human's perspective ($\mathcal{M}_R^*$). In other words, the robot needs to understand the model of itself in the human's eyes. Typically, model learning is addressed in the context of learning from demonstration [2], inverse reinforcement learning [1], and tutoring systems [9]. These approaches are not suitable for generating explicable behavior, since the robots need to learn the behavior that is expected of them by the humans, which may or may not reflect how the humans themselves would behave. As we shall see, in our work, $\mathcal{M}_R^*$ is learned through traces of the robot labeled by the humans in a particular way.

There exists work on generating legible motions [6] which considers a similar issue in motion planning. We are, on the other hand, concerned with task planning. Note that two different task plans may map to exactly the same motions which can be interpreted vastly differently by humans. In such cases, considering only motion becomes insufficient. There also exists work on the concept of cross training [10] which is to "cross-train" human-robot teams to reach a model consensus. Our work is concerned with learning model differences and how they influence the human's understanding of the robot behavior.

III. Plan Explicability

In our settings, an agent R needs to achieve a goal given by a human who is in the same environment. In this paper, we assume that the robot model M_R is based on PDDL [7]. As we discussed, for an agent to generate explicable plans, it must not only consider M_R but also $\mathcal{M}_R^*$. Given a domain, the problem is to find a plan for a given goal that satisfies:

$$\operatorname*{argmin}_{\pi_{M_R}} cost(\pi_{M_R}) + \alpha \cdot dist(\pi_{M_R}, \pi_{\mathcal{M}_R^*}) \tag{1}$$

where π_{M_R} is a plan that is constructed using M_R (i.e., the agent's plan), $\pi_{\mathcal{M}_R^*}$ is a plan that is constructed using $\mathcal{M}_R^*$ (i.e., the human's anticipation of the agent's plan), $cost$ returns the cost of a plan, $dist$ returns the distance (i.e., capturing the differences) between two plans, and α is the relative weight. The goal of Eq. (1) is to find a plan that minimizes a weighted sum of the cost of the agent plan and the differences between the two plans. Since the agent model M_R is assumed to be given, the challenge lies in the second part in Eq. (1).

If we know $\mathcal{M}_R^*$ or it can be learned, the only thing left would be to search for a proper $dist$ function. However, as discussed previously, $\mathcal{M}_R^*$ is inherently hidden, difficult to convey, and can be arbitrarily different from M_R. Hence, our solution is to use a learning method to directly approximate the returned values. We postulate that humans understand agent plans by associating abstract tasks with actions, which can be considered as a labeling process. Based on this, we assume that $dist(\pi_{M_R}, \pi_{\mathcal{M}_R^*})$ can be functionally decomposed as:

$$dist(\pi_{M_R}, \pi_{\mathcal{M}_R^*}) = F \circ \mathcal{L}^*(\pi_{M_R}) \tag{2}$$

where F is a domain specific function that takes plan labels as input, and $\mathcal{L}^*$ is the labeling scheme of the human for agent plans based on $\mathcal{M}_R^*$. As a result, Eq. (1) now becomes:

$$\operatorname*{argmin}_{\pi_{M_R}} cost(\pi_{M_R}) + \alpha \cdot F \circ \mathcal{L}_{CRF}^*(\pi_{M_R} | \{S_i | S_i = \mathcal{L}^*(\pi_{M_R}^i)\}) \tag{3}$$

where $\{S_i\}$ is the set of training examples and $\mathcal{L}_{CRF}^*$ is the learned model of $\mathcal{L}^*$. We can now formally define plan explicability in our context. Given a plan of agent R as a sequence of actions, π_{M_R} (simplified below as π for clarity):

$$\pi = \langle a_0, a_1, a_2, ...a_N \rangle \tag{4}$$

where a_0 is a null action that denotes plan starting. Given the domain, we assume that a set of task labels T is provided to label agent actions:

$$T = \{T_1, T_2, ...T_M\} \tag{5}$$

Explicability is concerned with the association between abstract tasks and agent actions; each action in a plan is associated with an action label. The set of action labels (L) for explicability is the power set of the task labels: $L = 2^T$. When an action label includes multiple task labels, the action is interpreted as contributing to multiple tasks; when an action label is the empty set, the action is interpreted as inexplicable. When a plan is labeled, we can compute its explicability measure based on its action labels in a domain specific way.

Definition 1 (Plan explicability): Given a domain, the explicability θ_π of an agent plan π is computed by a mapping, $F_\theta : \mathbf{L}_\pi \rightarrow [0, 1]$ (with 1 being the most explicable).

$\mathbf{L}_\pi$ above denotes the sequence of action labels for π. An example of F_θ used in our evaluation is given below:

$$F_\theta(\mathbf{L}_\pi) = \frac{\sum_{i \in [1,N]} \mathbf{1}_{L(a_i) \neq \emptyset}}{N} \tag{6}$$

where N is the plan length, $L(a_i)$ returns the action label of a_i, and $\mathbf{1}_{formula}$ is an indicator function that returns 1 when the $formula$ holds or 0 otherwise. Eq. (6) basically computes the ratio between the number of actions with non-empty action labels and the number of all actions. Please also refer to [15] for how this labeling process is used to learn the CRFs for label prediction for computing the explicability measure.

IV. Evaluation

Here, we evaluate our approach with human subjects using physical robots in a blocks world domain. In this domain, the robot's goal (which is known to the human) is to build a tower of a certain height using blocks on the table. The towers to be built have different heights in different problems. There are two types of blocks, light ones and heavy ones, which are indistinguishable externally but the robot can identify them based on the markers. Picking up the heavy blocks are more costly than the light blocks for the robot. Hence, the robot may sometimes choose seemingly more costly (i.e., longer) plans to build a tower from the human's perspective. We incorporate the explicability measure as a heuristic into the FastForward (FF) planner with enforced hill climbing [8] (for details see [15]). We evaluate plans generated by the robot using our planner (FF-EXPD) and a cost-optimal planner (OPT).

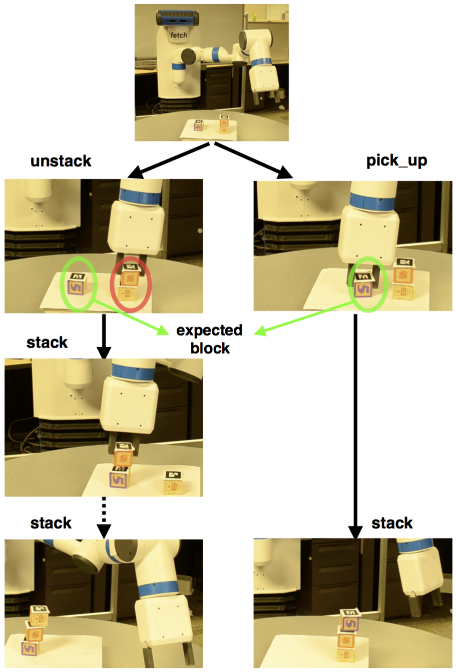

Fig. 2. Execution of two plans generated by OPT (left) and FF-EXPD (right) for one out of the 8 testing scenarios (`https://www.youtube.com/watch?v=AAAwSVbAV7s`). The top figure shows the setup where the goal is to build a tower of height 3. The block that is initially on the left side of the table is a heavy block. The optimal plan involves more actions with the light blocks (i.e., putting the two light blocks on top of the heavy one) while the explicable plan is more costly as it requires moving the heavy one.

We generated a set of 23 problems in this domain in which towers of height 3 are to be built. The plans for these problems were manually generated and labeled as the training set. For 4 out of these 23 problems, the optimal plan is not the most explicable plan. We then generated a set of 8 testing problems for building towers of various heights (from $3-5$) to verify that our approach can generalize. Testing problems were generated only for cases where plans are more likely to be inexplicable. For each problem, we generated two plans, one using OPT and the other using FF-EXPD, and recorded the execution of these plans on the robot. We recruited 13 subjects on campus and each human subject was tasked with labeling two plans (generated by OPT and FF-EXPD respectively) for each of the 8 testing problems, using the recorded videos. After labeling each plan, we also asked the subject to provide a score ($1-10$ with 10 being the most explicable) to describe how comprehensible the plan was overall.

In this evaluation, we only use one task label "*building tower*". For all testing problems, the labeling process results in 77.8% explicable actions (i.e., actions with a task label) for OPT and 97.3% explicable actions for FF-EXPD. The average explicability measures for FF-EXPD and OPT are 0.98 and 0.78, and the average scores are 9.65 and 6.92, respectively. We analyze the results using a paired T-test which shows a significant difference between FF-EXPD and OPT in terms of the explicability measure (using Eq. (6)) computed from the human labels and the overall scores ($p < 0.001$ for both). Furthermore, after normalizing the scores from the human subjects, the Cronbach's α value shows that the explicability measures and the scores are consistent for both FF-EXPD and OPT ($\alpha = 0.78, 0.67$, respectively). These results verify that: 1) our explicability measure does capture the human's interpretation of the robot plans and 2) our approach can generate plans that are more explicable to humans. In Fig. 2, we present the plans for a testing scenario.

V. Conclusion

In this paper, we introduced plan explicability to synthesize plans that are more comprehensible to humans. To achieve this, robots must consider not only their own models but also the human's interpretation of their models. This enables agents to synthesize plans that can be easily understood by humans. To the best of our knowledge, this is the first attempt to model plan explicability for task planning. The proposed measure has a variety of applications (e.g., achieving fluent human-robot interaction and ensuring human safety). To compute this measure, we learn the labeling scheme of humans for agent plans from training examples based on CRFs. We then use this learned model to label a new plan to predict its explicability.

Finally, while we focus on robot task planning, our work also has many other interesting applications. For example, many defense applications use planning to create inexplicable plans, which can help deter or confuse enemies and are also useful for testing defenses against new or unexpected attacks.

Acknowledgment: This research is supported in part by the ONR grants N00014-13-1-0176, N00014-13-1-0519 and N00014-15-1-2027, and ARO grant W911NF-13-1-0023.

References

[1] P. Abbeel and A.Y. Ng. Apprenticeship learning via inverse reinforcement learning. In *ICML*, 2004.

[2] B.D. Argall, S. Chernova, M. Veloso, and B. Browning. A survey of robot learning from demonstration. *Robot. Auton. Syst.*, 2009.

[3] T. Chakraborti, G. Briggs, K. Talamadupula, Y. Zhang, M. Scheutz, D. Smith, and S. Kambhampati. Planning for serendipity. In *IROS*, 2015.

[4] M. Cirillo, L. Karlsson, and A. Saffiotti. Human-aware task planning for mobile robots. In *ICAR*, pages 1–7, June 2009.

[5] G. Csibra and G. Gergely. 'obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta psychologica*, 124(1):60–78, 2007.

[6] A. Dragan and S. Srinivasa. Generating legible motion. In *RSS*, 2013.

[7] M. Fox and D. Long. PDDL2.1: An extension to PDDL for expressing temporal planning domains. *J. Artif. Int. Res.*, 20(1), December 2003.

[8] J. Hoffmann and B. Nebel. The FF planning system: Fast plan generation through heuristic search. *J. Artif. Int. Res.*, 14(1):253–302, May 2001.

[9] T. Murray. Authoring Intelligent Tutoring Systems: An analysis of the state of the art. *Intl. J. of Artif. Int. in Education*, 10:98–129, 1999.

[10] S. Nikolaidis, P. Lasota, R. Ramakrishnan, and J. Shah. Improved human-robot team performance through cross-training, an approach inspired by human team training practices. *IJRR*, 34(14), 2015.

[11] K. Talamadupula, G. Briggs, T. Chakraborti, M. Scheutz, and S. Kambhampati. Coordination in human-robot teams using mental modeling and plan recognition. In *IROS*, 2014.

[12] R.R. Vallacher and D.M. Wegner. What do people think they're doing? action identification and human behavior. *Psychological Review*, 1987.

[13] Y. Zhang, V. Narayanan, T. Chakraborty, and S. Kambhampati. A human factors analysis of proactive assistance in human-robot teaming. In *IROS*, 2015.

[14] Y. Zhang, S. Sreedharan, and S. Kambhampati. Capability models and their applications in planning. In *AAMAS*, 2015.

[15] Y. Zhang, S. Sreedharan, A. Kulkarni, T. Chakraborti, H. Hankui Zhuo, and S. Kambhampati. Plan Explicability and Predictability for Robot Task Planning. *arXiv: 1511.08158 [cs.AI]*, November 2015.