Explicable Plans for Human-Robot Teams

Anagha Kulkarni, Sarath Sreedharan, Yu Zhang, Tathagata Chakraborti and Subbarao Kambhampati

***Abstract*—For the robots to be part of our daily lives, their ability to form safe and successful collaborations with humans is a necessity. To ensure smoother collaborations, we provide a framework for human-robot teams using the concept of explicability. Explicability outlines that an agent's perception of another agent's model may be different from the actual model and that this should be factored into the agent's planning process. Our contributions include extending explicability formulation to support interactive human-robot teaming and implementing the framework on a physical robotic platform. We make a reasonable assumption that an agent might only have an approximate version of other agent's model, i.e. the robot not only has to learn human's preconceptions about its own model, but also has to work with incomplete human planning preferences.**

I. Introduction

As more and more robots are being placed in human society, it has become crucial to address the challenges of human-robot interaction. An important challenge is to ensure that the robot always behaves in expected and intuitive fashion when interacting with humans. Unexpected behaviors in robots can raise safety concerns. Especially in human-robot teams, to ensure co-ordination and good teamwork between the agents, generation of explicable behavior can be of utmost importance. For a team to achieve their common goal efficiently, it is necessary that the agent is aware of the other agents' individual goals. In a human-robot team, robot has to be mindful of human's tasks, and avoid hampering human's personal goals while achieving its own goals. In essence, there has to be a mutual understanding between the participating agents regarding what actions to take at every step.

The notion of human-robot teaming has been a popular research direction. There have been many works involving human-robot teams with robots as proactive agents [9], [10], [1]. In most such works, the assumption is that the human model is provided and complete for inferring about the human intent and plan. This is often not true. However, [11], presents a new outlook on human-robot teaming, where the robot also factors in human's mental model of the robot's model in its decision making, in order to generate plans that are more explicable and comprehensible to humans. This work is closely related to "legible" motion in motion planning [5] and generating socially acceptable behaviors for robots [7], [6]. While our work is inspired by [11], we extend the framework to support interactive human-robot teaming instead of the human just being an observer as in [11]. We extend the formulation of explicablity to capture other behaviors that may be important for successful teaming like altruism [2], opportunism [3], compliance [2], etc. To this end, we learn a model that is able to capture both human mental model of the robot and to a limited degree the human planning preferences. We then evaluate our model using a Fetch robot on a *blocksworld* domain, where human and robot collaborate to form meaningful words out of lettered blocks.

II. Problem Formulation

We build on the formulation given in [11], please refer section III of that paper. Here we consider two member peer-to-peer human-robot teams. Here each problem can be written as a planning problem $P_T = \langle I, M_R, \widetilde{M_H}, \Pi_C, G_H, G_R \rangle$, where I denotes the initial state of the planning problem, G_H and G_R represent the goals of human and robot, M_R is the actual robot model, $\widetilde{M_H}$ is the approximate human planning model provided to the robot and Π_C represents a set of annotated plans that are used to train the CRF model. The plan for the entire team will be represented through a composite plan. A composite plan is defined as follows:

Definition 1. *A composite plan π_C captures the actions performed by both human and robot to achieve their goals (common goals and individual) and is represented as $\pi_C = \{a_1^{\phi_1}, a_2^{\phi_2}, ..., a_i^{\phi_i} ..., a_n^{\phi_n}\}$. Here $a_i^{\phi_i}$ represents the i^{th} action in the plan performed by the agent ϕ_i (ϕ_i can be H or R).*

Here we assume that only one agent is executing its action at any given time and that the actions are performed alternately i.e if $\phi_i = H$ then $\phi_{i+1} = R$. We allow a *noop* action, which can be used any time an agent wants to skip its turn. The robot has to generate a composite plan that is close to the plan that the human might generate. It does this by embedding its actions in a composite plan that includes its expectations of the actions the human is likely to take:

$$\underset{\pi_C^{M_R,\widetilde{M_H}}}{\arg\min}\, cost(\pi_C^{M_R,\widetilde{M_H}}) + \alpha.dist(\pi_C^{M_R,\widetilde{M_H}}, \pi_C^{\widetilde{M_R},M_H}) \quad (1)$$

where $\pi_C^{M_R,\widetilde{M_H}}$ is the composite plan created by the robot using M_R and $\widetilde{M_H}$, while $\pi_C^{\widetilde{M_R},M_H}$ is the composite plan that might be created by the human. As in [11], the distance function $dist(\pi_C^{M_R,\widetilde{M_H}}, \pi_C^{\widetilde{M_R},M_H})$ can be calculated as a function of labels of actions in $\pi_C^{M_R,\widetilde{M_H}}$ which gives us equation 2:

$$\begin{gathered}\underset{\pi_C^{M_R,\widetilde{M_H}}}{\arg\min}\, cost(\pi_C^{M_R,\widetilde{M_H}}) + \\ \alpha.F \circ L^*_{CRF}(\pi_C^{M_R,\widetilde{M_H}} | \{S_i | S_i = L^*(\pi_C^{i*})\})\end{gathered} \quad (2)$$

As shown in equation 2, the label for each action is produced by a CRF (L^*_{CRF}) trained on a set of labeled execution traces ($\{\pi_C^i*\}$). Each action was labeled by human subjects and each label was chosen from $T = \{t_1^H, t_2^R, ..., t_m^R\}$, where T is the set of all high level sub-goals. Equation 2 represents the objective of the problem, where the $g\ value$ is given by the term $cost(\pi_C^{M_R,\widetilde{M_H}})$ and $h\ value$ is given by the distance function (i.e $\alpha.F \circ L^*_{CRF}(\pi_C^{M_R,\widetilde{M_H}}|\{S_i|S_i = L^*(\pi_C^{i*})\})$).

In addition to capturing whether humans can relate a specific action to a specific subgoal, we also try to reason about usefulness of sub-goal to team goal. To do this, we introduce a set of higher level semantic class labels S, where $S = \{S_1, S_2, \ldots, S_k\}$ and we have $\forall t \in T, \exists S_i \in S,$ such that $t \in S_i$. For example, an action like *Pickup A* could correspond to a sub-goal *Fetch next block*, which in turn could fall under a class label like *Robot working on its goal*. By introducing these semantic class labels, we extend explicability formulation to capture behaviors like altruism, opportunism, compliance, etc. We use explicability score to generate our heuristic for the planning. We compute this score by assigning different weights for each class label. The exact weights of each class label depends on the desired behavior. For example, if altruistic behavior is desired, we can set high value for class label that comprises of actions where robot is helping human in her goal, else if opportunistic behavior is desired we can set high value for class label that comprises actions where human helps robot.

III. Evaluation

To evaluate our system, we test it on modified blocksworld domain, with lettered blocks. The goal was to form different words by stacking the blocks with the required letters. For each scenario, both agents were assigned unique words, and each agent was aware of other's goal. We used Fetch robot as the robot agent. The entire workspace is divided into two sections near and far, where accessing the blocks at farther region incurs some extra cost. The domain model for each agent was based on the IPC blocksworld domain with four standard actions ($Stack$, $Unstack$, $Pickup$ and $Putdown$) and an additional $noop$ action which had no precondition and effects. All training examples were collected from human subjects, with random initial and goal states. We collected an initial set of 70 traces, which was used to produce more unique plan traces by substituting the blocks-IDs, changing the block letters and final goal. After this process we had a total of 14,000 traces, which were used to train CRF model.

Use case: Consider a scenario (demo video[1]), where the initial state is as shown in the Figure 1. The robot's goal is to form the word EAT, while the human has to form PEN. The initial state has two E's such that one E is near and another is farther away. Here we consider two plans: one optimal plan, another explicable plan generated by modified FF planner [8] sensitive to explicability. With altruistic behavior desired from the robot, we test the explicability score of both the plans.

Here the robot should use farther away E to complete its goal, or it should bring the farther E to nearer region for the sake of the human. In essence, the robot should not inconvenience the human by leaving farther block for her. The optimal plan involves robot picking the closer block to reduce its plan cost. On the other hand, in the explicable plan the

[1]https://www.dropbox.com/s/bl6fyrlp052h0za/EXP-VIDEO.mp4?dl=0

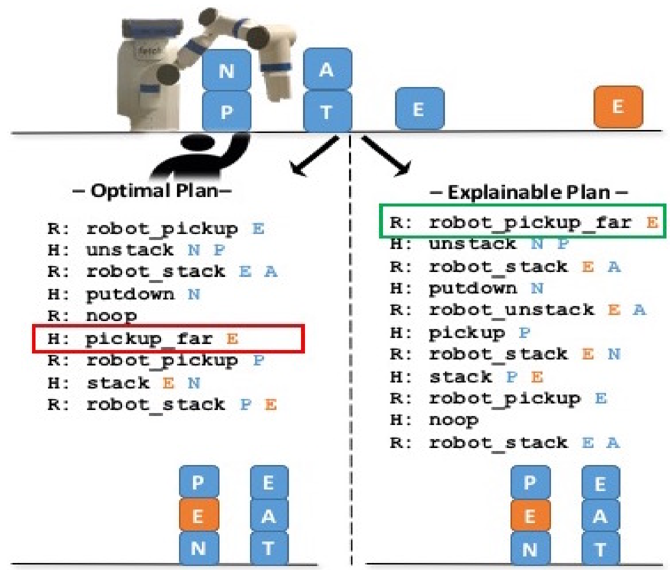

Fig. 1. In first case human picks farther E, in the other robot picks it

robot helps the human by picking the farther E and then further helps the human by stacking E on human's blocks. We use trained CRF model to get class labels for actions. If class labels represent inexplicable or disruptive actions on part of the robot, weight assigned is -3, if agent is working on it's own goal, weight is 1, if an agent is helping other agent complete its goal, weight is 3. For the optimal plan, the score is 0.111, whereas for explicable plan it is 0.818.

IV. Discussion

Although we have mainly focused on two member teams, we believe that this framework can be easily extended to larger team sizes. One of the main challenges in larger team sizes would be to maintain the order in which agents may choose to perform actions. Another assumption was that all action executions were sequential, it would be interesting to see if this formulation can be extended to support simultaneous action executions. One way to achieve this would be by using temporal planners [4]. In conclusion this work aims at introducing a way of creating plans for human robot teams, that are naturally more explicable and preferred by the humans.

References

[1] Tathagata Chakraborti, Gordon Briggs, Kartik Talamadupula, Yu Zhang, Matthias Scheutz, David Smith, and Subbarao Kambhampati. Planning for serendipity. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 5300–5306. IEEE, 2015.

[2] Tathagata Chakraborti, Kartik Talamadupula, Yu Zhang, and Subbarao Kambhampati. Interaction in human-robot societies. In *AAAI Workshop on Symbiotic Cognitive Systems*, 2016.

[3] Tathagata Chakraborti, Yu Zhang, David Smith, and Subbarao Kambhampati. Planning with resource conflicts in human-robot cohabitation. In *AAMAS*, 2016.

[4] Minh B. Do and Subbarao Kambhampati. Sapa: A multi-objective metric temporal planner. *J. Artif. Int. Res.*, 20(1): 155–194, December 2003. ISSN 1076-9757. URL http://dl.acm.org/citation.cfm?id=1622452.1622460.

[5] Anca Dragan and Siddhartha Srinivasa. Generating legible motion. In *Proceedings of Robotics: Science and Systems*, Berlin, Germany, June 2013.

[6] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. A survey of socially interactive robots. *Robotics and autonomous systems*, 42(3):143–166, 2003.

[7] Guy Hoffman and Cynthia Breazeal. Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, pages 1–8. ACM, 2007.

[8] Jörg Hoffmann and Bernhard Nebel. The ff planning system: Fast plan generation through heuristic search. *Journal of Artificial Intelligence Research*, pages 253–302, 2001.

[9] Julie Shah, James Wiken, Brian Williams, and Cynthia Breazeal. Improved human-robot team performance using chaski, a human-inspired plan execution system. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 29–36. ACM, 2011.

[10] Yu Zhang, Vignesh Narayanan, Tathagata Chakraborty, and Subbarao Kambhampati. A human factors analysis of proactive assistance in human-robot teaming. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2015.

[11] Yu Zhang, Sarath Sreedharan, Anagha Kulkarni, Tathagata Chakraborti, Hankz Hankui Zhuo, and Subbarao Kambhampati. Plan explicability for robot task planning. In *Proceedings of the RSS Workshop on Planning for Human-Robot Interaction: Shared Autonomy and Collaborative Robotics*, 2016.