Generating Active Explicable Plans in Human-Robot Teaming

Akkamahadevi Hanni and Yu Zhang

***Abstract*— Intelligent robots are redefining a multitude of critical domains but are still far from being fully capable of assisting human peers in day-to-day tasks. An important requirement of collaboration is for each teammate to maintain and respect an understanding of the others' expectations of itself. Lack of which may lead to serious issues such as loose coordination between teammates, reduced situation awareness, and ultimately teaming failures. Hence, it is important for robots to behave explicably by meeting the human's expectations. One of the challenges here is that the expectations of the human are often hidden and can change dynamically as the human interacts with the robot. However, existing approaches to generating explicable plans often assume that the human's expectations are known and static. In this paper, we propose the idea of active explicable planning to relax this assumption. We apply a Bayesian approach to model and predict dynamic human belief and expectations to make explicable planning more anticipatory. We hypothesize that active explicable plans can be more efficient and explicable at the same time, when compared to explicable plans generated by the existing methods. In our experimental evaluation, we verify that our approach generates more efficient explicable plans while successfully capturing the dynamic belief change of the human teammate.**

I. Introduction

Advancements in robotics has opened up many opportunities for novel robotic applications. For example, robots can play a vital role in helping humans carry out everyday tasks such as house chores and accomplishing critical missions like urban search and rescue. In many cases, it is desirable for the robots to be fully autonomous and minimize human intervention. As a result, a desirable capability is to have these robots maintain awareness of human presence and their expectations of the robot in the environment. Absence of this capability may lead to serious consequences such as reduced team situation awareness and loss of trust.

The authors in [1] first proposed to incorporate the human's expectation in the robot's decision-making. One of the key assumptions there is that humans generate their expectations of the robot based on a model of expectation in the human's mind, much like how the robot generates behaviors using its domain model. As a result, the robot should make an effort to align its behavior with the human's expectation. A similar idea is adopted in other work, such as in [2] and [3]. However, these prior approaches all assume that the model for generating human expectations remains fixed throughout the task. While such an assumption maybe innocent to make in scenarios where the entire plan is presented to the human at once, it will definitely fail to hold in human-robot teaming scenarios where the human observes the robot's actions and revise the expectations of the robot over time. For example, if a human saw a robot fly, the robot would be expected to fly afterwards when needed, even if the human did not know about it in the first place.

The authors are with the School of Computing, Informatics, and Decision Systems Engineering at Arizona State University, Tempe, Arizona, USA. {ahanni, yzhan442}@asu.edu.

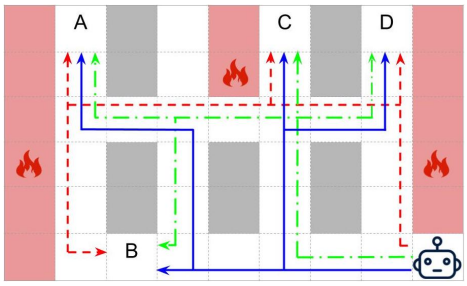


Fig. 1: USAR domain: building layout with the paths that correspond to the optimal plan - OP (Green), explicable plan - EXP (Red) and active explicable plan - ActiveEXP (Blue).

OP (Green)	EXP (Red)	ActiveEXP (Blue)
Collect C	Collect D	Collect C
Collect D	Collect C	Collect D
Collect A	Collect A	Collect B
Collect B	Collect B	Collect A

TABLE I: Steps of the three different plans in Fig. 1

In this paper, we introduce the notion of active explicable planning[1] to relax this assumption. We formulate the problem in a Bayesian framework by modeling changes of the human's model and expectations in a Dynamic Bayesian Network. We then derive a new explicable planning solution based on the changes anticipated in the human's expectations throughout the plan, assuming that the human always observes the robot's actions. In our experiments, we show that the plan generated by our approach is simultaneously more efficient and explicable. Results also confirm that humans do update their model and expectations dynamically.

The contribution of our work is three-fold: first, we introduce active explicable planning to relax a restrictive assumption of the existing work. Second, we derive a novel Bayesian formulation for active explicable planning and propose a solution. Third, we evaluate the proposed method to demonstrate its effectiveness in human-robot teaming.

A. Motivational Example

Consider yourself working with a robot in an urban search and rescue scenario, which involves retrieving valuables (A,

[1]This work was first presented as a Late-Breaking Report at HRI 2021. Here, we significantly extend the work to include the derivation of the Bayesian formulation and solution method, as well as a more comprehensive result section that includes both synthetic and human subject evaluations.

B, C and D) from a building under fire (see Fig. 1). While the robot's task is to retrieve valuables from the building safely, your task is to supervise the robot. Assume you have information about the internal layout of the building, locations of the valuables, and areas in the building on fire. The plan you would expect the robot to execute may be the shortest-path plan, i.e., the one shown in Red (Collect D, Collect C, Collect A and Collect B). This is shown also as the EXP plan in Table I. However, the robot could be sensitive to high temperatures and would prefer to avoid passing through *long corridors* near a fire. In such cases, the optimal plan for the robot would be the path indicated in Green (Collect C, Collect D, Collect A and Collect B), shown as the OP plan.

Now consider another plan indicated in Blue (Collect C, Collect D, Collect B and Collect A), the ActiveEXP plan. When the robot first collects C, it is likely that you would consider that the robot is trying to avoid the corridor near the fire. You would then consider navigating near a fire as undesirable. As such, the behavior of the robot to avoid the fire passage near C by moving around it to collect B becomes expected. Notice that even though the optimal plan also avoids the long corridor towards D and collects C first, it also crosses the short passage near a fire (near C) to collect A, which could cause confusion. A key observation here is that the ActiveEXP plan incurs slightly more cost than the optimal plan due to the detour while still being explicable due to the dynamic nature of the human's understanding. Without modeling the dynamic change of the human's understanding, it would be difficult to generate plans like ActiveEXP.

II. Related Work

Explicable plan generation falls under the umbrella of explainable planning [4], which has attracted a significant amount of attention. Methods have been proposed under various similar but different notations that include explicability [1], [5], [2], interpretability [6], [7], predictability [6], legibility [8], [3], and transparency [9]. One of the common features of such methods is the focus on assisting the understanding of different aspects of the plan, whether it is about the goal, actions, or rationale of the plan.

Our work on active explicable planning is closely connected to plan explicability [1], [2], where the aim is to generate behaviors that are expected by the human teammate, assuming that the goal is known. This is critical in human-robot teaming tasks since any inconsistency between the robot's behavior and human's expectation can lead to reduced situation awareness and the loss of trust. In [2], the authors assume that the human's model of expectation is known. It then learns a regression model to predict a distance metric between the robot's plan and the expected plan, which is used to choose a plan that is closer to the expected plan. Authors in [1] relax the assumption of a known human model. They learn the human's expectations via a sequential labeling scheme. A more principled way to relax this assumption is to maintain a belief of the model of expectation, similar to [7]. However, these prior works all assume that the human's model of expectation remains fixed throughout the plan.

The dynamic modeling of the human's understanding of the robot in our work has close connections to intent, plan and goal recognition [10], [11], [12], [13], where observations inform the recognition. It more generally involves the dynamic modeling of mental state, such as trust [14], [15], [16], except that the mental state to be modeled here is the human's understanding of the robot's domain model (i.e., the model of expectation). While it may be tempting to apply a POMDP framework [17], it is impractical as the observation function and the explicability score (to be discussed soon) in our problem are plan-context dependent, and hence the POMDP will be computationally expensive to solve.

III. Problem Formulation

In this work, we represent domain models using the PDDL language [18]. A model $\mathcal{M}$ = (($\mathcal{F}$, $\mathcal{A}$), $\mathcal{I}$, $\mathcal{G}$) where $\mathcal{F}$ and $\mathcal{A}$ are a finite set of fluents and actions, respectively. Also, $\forall a \in \mathcal{A}$, an action a is specified as a set of preconditions Pre(a), add effects Add(a), and delete effect Del(a), which are subsets of $\mathcal{F}$. Each action is also associated with a cost c(a). Let S be the set of states where $s \in S$ is a unique instantiation of the set of fluents $\mathcal{F}$. The initial state $\mathcal{I} \in S$ and the goal state $\mathcal{G} \in S$. We assume that the initial state and goal states are known to the human and robot and both the human and robot are rational agents. We first define a list of terms used throughout the paper:

- M_R is the model for generating the robot's behavior.
- M_H is the human's model of expectation (i.e., an understanding of the robot's model M_R).
- b_0 (M_H) is the initial belief of the human's understanding of the robot.
- I is the initial state.
- G is the goal state.
- c(π) is cost of the plan π
- γ is a weighting factor.

Definition 3.1: An **explicable planning** problem [2], [6] is defined as: given a tuple $\mathcal{P}_\mathcal{E}$ = $\langle$ M_R, M_H, I, G $\rangle$, find a plan $\pi_\mathcal{E}$ that satisfies the following:

$$\pi_\mathcal{E} = \underset{\pi_{M_R}}{\operatorname{argmin}} \ \phi(\pi_{M_R}, \pi_{M_H}) + \gamma c(\pi_{M_R}) \tag{1}$$

where, ϕ is the explicability distance that captures the difference between the plans π_{M_R} and π_{M_H}, which are generated by M_R and M_H, respectively.

To extend the formulation above to a dynamic human model, we define the problem of active explicable planning.

Definition 3.2: An **active explicable planning** problem is defined as: given a tuple $\mathcal{P}_{\mathcal{E}_\mathcal{A}}$ = $\langle$ M_R, b_0 (M_H), I, G $\rangle$, find a plan that satisfies the following:

$$\pi_{\mathcal{E}_\mathcal{A}} = \underset{\pi_{M_R}}{\operatorname{argmax}} \ \mathcal{E}_\mathcal{A}(\pi_{M_R}) - \gamma \, c(\pi_{M_R}) \tag{2}$$

where, $\mathcal{E}_\mathcal{A}$ is the active explicability score of a plan (which we will explain in the next section).

Again, the two main differences between an explicable planning problem $\mathcal{P}_\mathcal{E}$ and an active explicable planning problem $\mathcal{P}_{\mathcal{E}_\mathcal{A}}$ are: 1) $\mathcal{P}_\mathcal{E}$ assumes a fixed human model whereas

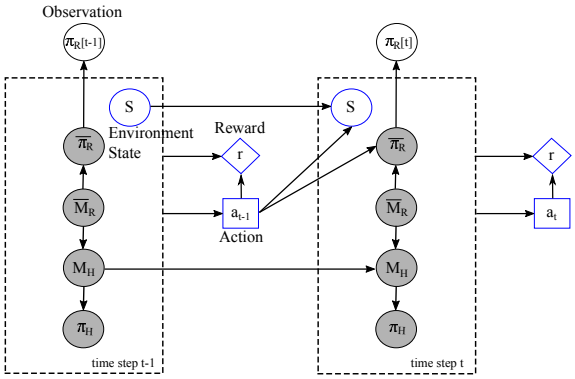


Fig. 2: Influence Diagram for active explicable planning for modeling dynamic human belief M_H. In this model, we assume that the human uses M_H to generate expectations of the robot's behavior. Furthermore, the human also maintains a temporary robot model $\overline{M}_R$ (different from M_R) to capture the changes to be made to the human's model. These changes are introduced as a result of observing the robot's current action, when it is contrasted to the plan introduced by $\overline{M}_R$.

$\mathcal{P}_{\mathcal{E}_\mathcal{A}}$ considers that this model may change dynamically during a plan execution, and 2) $\mathcal{P}_{\mathcal{E}_\mathcal{A}}$ operates on a belief or a distribution of possible models.

To address this problem, we assume that the human is a Bayesian rational (noisily rational) [19] thinker. This allows us to model how the human's understanding of the robot may change as observations are made about the robot. We also assume that the model space M is known and both M_R and M_H belong to M. Next, we introduce a few notions that will be used later. Given a pair $\{I,\ G\}$, a candidate plan space $\Pi_i = \{\pi_1, \pi_2, ..., \pi_p\}$ is a finite set of p plans generated by a possible model $M_i \in M$. All plans in Π_i are bounded by a cost threshold $\zeta \cdot \mathrm{c}(\pi^*)$, where $\zeta \geqslant 1$ and $\mathrm{c}(\pi^*)$ is the cost of the optimal plan for $\{I, G\}$ in M_i. The factor ζ can be arbitrarily set. In our evaluation, we set the cost threshold to be slightly above 1.

IV. Methodology

In explicable planning, the objective is to trade off the plan cost with explicability. In this work, we extend explicable planning to consider how the change in human's belief influences the expectations, as observations are made about the robot's actions.

A. Active Explicability

Instead of relying on a distance metric to quantify plan explicability as in [1], [2], [6], we adopt a Bayesian formulation, similar to some plan interpretibility measures in [7]. This allows us to define plan explicability measure in a principled way by directly comparing plans. However, the limitation is that it will not work when the human's and the robot's model spaces are different. We will delay further discussions on this to future work. More specifically, the explicability of a plan π_R in our work is defined as the probability of the plan π_R being in a set of candidate plans that can be generated by the human's model M_H. Intuitively, a plan that is more likely to be generated by the human's model is more explicable. The explicability of a plan π_R is given by:

$$\mathcal{E}(\pi_R) = f(\pi_R, M_H) = P(\pi_R \in \Pi_H | M_H) \tag{3}$$

where Π_H is the set of candidate plans in M_H with cost less than $\zeta \cdot \mathrm{c}(\pi^*)$. We further assume the probability of such plans being considered follows a noisy distribution (e.g, Boltzmann) based on the cost difference with π^*. When M_H is unknown, we consider a belief distribution over all possible models in M, represented by $b(M_H)$. The explicability of a plan π_R is defined as follows:

$$\begin{aligned}\mathcal{E}(\pi_R) &= f(\pi_R, b(M_H)) \\ &= \sum_{M_H} P(\pi_R \in \Pi_H | M_H) \cdot b(M_H)\end{aligned} \tag{4}$$

Furthermore, to capture that $b(M_H)$ can change over time, we define the active explicability of π_R to be the average of action explicability measures at each plan step as in [1]:

$$\begin{aligned}\mathcal{E}_\mathcal{A}(\pi_R) &= \frac{1}{T}\sum_t f(\pi_R[t], b^t(M_H^t)) \\ &= \frac{1}{T}\sum_t \sum_{M_H^t} P(\pi_R[t] \in \Pi_H^t | M_H^t) \cdot b^t(M_H^t)\end{aligned} \tag{5}$$

where $\pi_R[t]$ represents the action at step t in π_R and Π_H^t represents the set of candidate plans in M_H^t generated at time t with initial state S^t (i.e., the resulting state from I after executing the first $t-1$ actions in π_R) and goal state G. Notice that since new plans are generated at every time step with initial state as S^t, the cost of the optimal plan also changes respectively. This assumes that the human will change his/her expectation of the robot's plan based on the current state after observing some robot actions. In such a case, if the robot's next action matches with the first action of an expected plan from the current state, the action is considered to be consistent with that plan. Note that we slightly abuse the notation here and use $\in$ to check this consistency. To further clarify the notation, we incorporate a delta function $\delta(\pi_R[t] \in \pi_H^t)$ to indicate whether an observation $\pi_R[t]$ matches the first action of the candidate plan in Π_H^t. Consequently, we rewrite the equation above as:

$$\mathcal{E}_\mathcal{A}(\pi_R) = \frac{1}{T}\sum_t \sum_{M_H^t} \sum_{\pi_H^t} \delta(\pi_R[t] \in \pi_H^t) P(\pi_H^t | M_H^t) \cdot b^t(M_H^t) \tag{6}$$

where, $P(\pi_H^t | M_H^t)$ is the likelihood function described by a Boltzmann distribution based on cost difference given by:

$$P(\pi_H^t | M_H^t) \propto \epsilon^{-\beta \times \{c(\pi_H^t) - c(\pi^*)\}} \tag{7}$$

In theory, a plan is the most explicable when it maximizes the explicability score without considering the cost. However,

in practice, it is desirable to choose a plan that balances the explicability score with the cost of a plan [1]. Similarly, an active explicable plan $\pi_{\mathcal{E}_\mathcal{A}}$ is defined as a plan that maximizes a weighted sum of active explicability score and plan cost as described in Eq. 2.

B. Human Belief Update

The key to computing Eq. 6 and in turn Eq. 2 is $b^t(M_H^t)$. To simplify the problem, we assume that the environment state and robot action are observable and that the robot actions are deterministic. $\pi_R = \pi_R[1], ..\pi_R[t]..$ is the sequence of actions that are observed by the human at the respective time step.

Consider the dynamic belief network in Fig. 2. There are two factors that influence the human's belief when an observation is made at time t: the prior belief of the human model M_H^{t-1} and the belief of a temporary model of the robot $\overline{M}_R^t$ (which is also used to derive the robot's behavior) upon observing an action $\pi_R[t]$. The idea behind considering this temporary model variable is to capture the models that are consistent with the observation made, which inform updates to the human's belief. For example, consider the human expects the robot to be only capable of walking but at some point observes the robot fly. In such a case, $\overline{M}_R^t$ would be consistent with all the candidate models where the robot can fly.

The belief of the human model at a given time t, $b^t(M_H^t)$ can be computed using the Forward algorithm derived below:

$$b^t(M_H^t) = P(M_H^t|\pi_R[1:t]) \tag{8}$$

Given the graphical model in Fig. 2:

$$b^t(M_H^t) = \sum_{M_H^{t-1}} \sum_{\overline{M}_R^t} \sum_{\bar{\pi}_R^t} \sum_{\pi_H^t} P(M_H^t, M_H^{t-1}, \overline{M}_R^t, \bar{\pi}_R^t, \pi_H^t|\pi_R[t], \pi_R[1:t-1]) \tag{9}$$

Expand the above equation using the chain rule and simplify using conditional independencies we get:

$$b^t(M_H^t) = \sum_{M_H^{t-1}} P(M_H^{t-1}|\pi_R[1:t-1]) \sum_{\overline{M}_R^t} P(\overline{M}_R^t) P(M_H^t|M_H^{t-1}, \overline{M}_R^t) \sum_{\bar{\pi}_R^t} P(\bar{\pi}_R^t|\overline{M}_R^t)\delta(\pi_R[t] \in \bar{\pi}_R^t) \sum_{\pi_H^t} P(\pi_H^t|M_H^t) \tag{10}$$

where, $P(M_H^{t-1} \mid \pi_R[1:t-1])$ is nothing but $b^{t-1}(M_H^{t-1})$. The term $\sum_{\pi_H^t} P(\pi_H^t|M_H^t)$ sums to 1 so it does not influence the belief update. Thus,

$$b^t(M_H^t) = \sum_{M_H^{t-1}} b^{t-1}(M_H^{t-1}) \sum_{\overline{M}_R^t} P(\overline{M}_R^t) P(M_H^t|M_H^{t-1}, \overline{M}_R^t) \sum_{\bar{\pi}_R^t} P(\bar{\pi}_R^t|\overline{M}_R^t)\delta(\pi_R[t] \in \bar{\pi}_R^t) \tag{11}$$

where the likelihood function of $\bar{\pi}_R^t$ is similar to that of π_H^t described previously.

V. Evaluation

In order to evaluate our approach, we run experiments on the IPC Blocksworld domain. The objective here is to demonstrate how dynamically tracking the model can benefit the plan efficiency by reducing the trade-off for explicability. To show that this approach is actually effective, we further validate it with the help of human subjects in a Taxi domain. The objective is to test whether the plan generated by our approach is indeed viewed as explicable and preferred, compared to the baseline methods.

For a given domain, we generate the model space M by identifying features that specify the true model M_R. These features represent different aspects of the model that the human teammate could have a misunderstanding about (e.g. not know about). If the number of features identified is k, then the model space M is a power set of 2^k models. The plan space Π is the union of all plans generated by each of the models bounded by $\zeta \cdot c(\pi^*)$.

Dynamic inference: We create a Dynamic Bayesian Network with nodes and edges as described in the Fig. 2. We initialize M_H as a uniform distribution of all models in M. $\overline{M}_R$ being the human's belief of the temporary robot's model is updated when an observation is made, informing the updates to the human's belief M_H as a result of the influence from the observation. We solve the problem by unrolling the resulting DBN and use Forward inference algorithm to infer the belief updates after sequential observations. We compute the active explicability score of a plan π_R using Eq. 6.

For synthetic data, we perform the following evaluations: (1) When the human's belief M_H is carried forward from one task to another, the belief update process in Eq. 11 will converge to the true robot model M_R. (2) Robustness of the process against noisy observations. (3) Efficiency of the approach by comparing the Active Explicability score and the cost of the plans generated. For the human study we aim to validate the following two hypotheses:

$H1$ = *"Human's belief can dynamically change throughout a plan."*

$H2$ = *"The active explicable plan is comparable to explicable plans while being more cost-efficient."*

A. Evaluation on Synthetic domain

We use the Blocksworld domain to evaluate how well the dynamic belief update works. We identify $k = 4$ features in this domain that could potentially be the factors causing the mismatches between the human's belief and the robot's true model. As a result, $2^4 = 16$ possible models are obtained. To test (1), we run our approach on 10 different problems with different initial I and goal G states and record the belief update for each of the 16 models. For the first problem, the human's belief is initialized to be a uniform distribution while for the subsequent problems it is initialized with the updated human's belief from the previous problem. We do so in support of the intuition that the human carries his/her belief from one problem to another. From the graph in Fig. 3. (a), we observe a gradual rise in the belief of the robot's true model M_R starting from a uniform distribution. The belief

update is always consistent with the true model albeit being a bit slow. This is due to the observed action being generated by at least one plan in every model due to high similarities among the models in M.

Our approach heavily depends on the observations that could affect the dynamic belief update towards the true robot model. In order to investigate this, we measured the belief update for the true robot's model M_R for different noise levels between 0% to 40%. We found that the change in belief update for the true robot's model M_R dampens as the noise level increases. From Fig. 3. (b) we can observe this change. As expected, we found that even with higher noise levels the model updates towards the true robot model.

We evaluate (3) by testing our approach on four different problems in the Blocksworld domain. Table II shows the costs and active explicability scores of the plans respectively. We can observe that the ActiveEXP plans receive higher $\mathcal{E}_\mathcal{A}$ scores than OP and EXP plans. In these samples, the ActiveEXP plans overlap with the OP plans in almost all problems due to the limited number of plans generated within the threshold. However, this is not always the case in general.

	OP		EXP		ActiveEXP	
Problem	cost	$\mathcal{E}_\mathcal{A}$	cost	$\mathcal{E}_\mathcal{A}$	cost	$\mathcal{E}_\mathcal{A}$
1	4	0.637	6	0.509	4	0.637
2	6	0.388	7	0.375	6	0.437
3	8	0.562	10	0.45	8	0.562
4	8	0.674	9	0.6	8	0.674

TABLE II: Cost and $\mathcal{E}_\mathcal{A}$ scores for OP, EXP and ActiveEXP plans generated for different problems in the Blocksworld.

B. Evaluation using Human Study

We design a Taxi domain to evaluate the effectiveness of our approach with human subjects. In the Taxi domain, a private taxi is required to pick up four guests among {A, B, C, D, E, F, G, H} from different locations and drop off at the Convention Center (CC). The subjects are provided with a GPS map with locations of the guests and the taxi at the CC. The subjects are informed that picking up guests that are farther (marked in Red) pays more money than those who are closer (marked in Green). However, the traffic situation (unknown to the subjects) can influence the taxi's decision to travel far away. We also inform the subjects that the taxi has access to the current traffic information that may influence its decision. We induce an initial bias of *light traffic* by informing the subjects that the current time is early morning hours. The traffic situation (Light/Heavy) is the hidden belief which we capture by querying the subjects at every step after revealing the taxi's schedule sequentially. The subject's task is to monitor the taxi agent's schedule. However, the real traffic situation is *heavy traffic*.

In this study, our goal is to compare the ActiveExp plan, the EXP plan and the Optimal plan with the help of human subjects via an Mturk study. We recruited 60 participants (20 for each plan). Before the human subjects made any observations, we asked them about their initial belief of the traffic situation to confirm our induced bias, which is also used as the prior belief. Then, we demonstrated the schedule of the robot sequentially by revealing the next guest to pick up and drop off. We also asked the subjects about their belief of the traffic situation at each step.

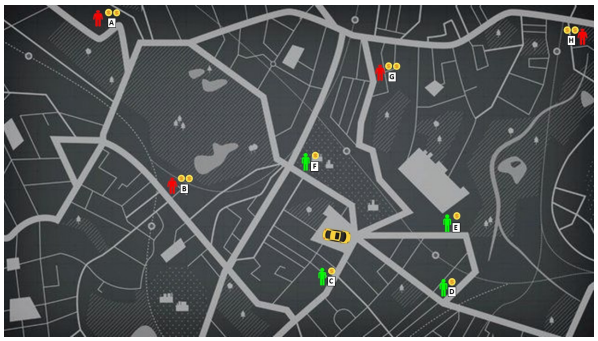


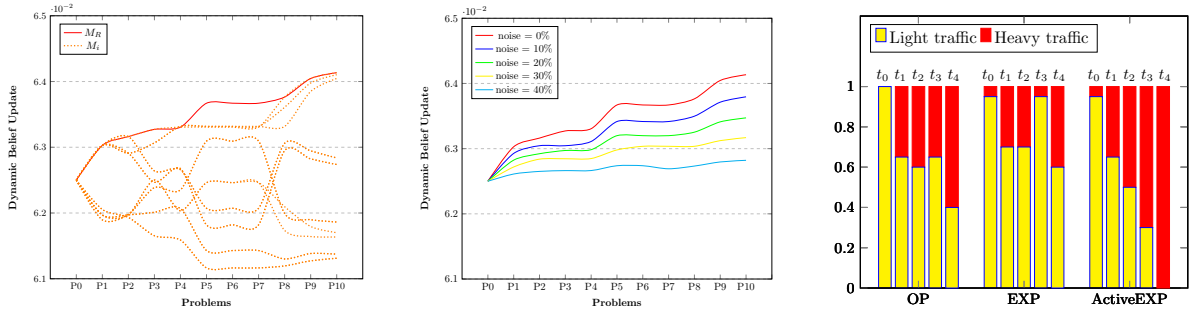
Fig. 4: The Taxi domain for human study

	OP	EXP	ActiveEXP
Plan	Pick & Drop G Pick & Drop C Pick & Drop B Pick & Drop F	Pick & Drop G Pick & Drop B Pick & Drop A Pick & Drop H	Pick & Drop G Pick & Drop C Pick & Drop D Pick & Drop E
Payoff	5.2	-2.4	4.2
$\mathcal{E}_\mathcal{A}$	0.123	0.161	0.143

TABLE III: Illustrative Plans OP, EXP and ActiveEXP. Guests marked in red (A, B, G and H) are farther away and pay double than the guests marked in Green (C, D, E and F) who are closer to the Convention Center.

Using this domain, we evaluate three plans ActiveEXP, EXP and OP shown in Table III to test $H1$ and $H2$. From Fig. 3. (c), we observe that the human belief changes dynamically from light traffic (initial belief) to heavy traffic after observing the taxi's schedule.

The optimal plan (OP) includes picking up G and B who are relatively closer than other guests marked in red, and C and F marked in green. This plan has the optimal payoff but evaluates to the lowest active explicability score ($\mathcal{E}_\mathcal{A}$). Intuitively, this plan does not convey the traffic situation to the observer well since it seems to have exhibited "oscillatory" decisions that point to ambiguous conclusions. Indeed the OP plan in Fig. 3. (c) shows mixed responses from subjects about their belief of the traffic situation. The explicable plan (EXP) includes picking up all guests marked in red (G, B, A and H). This plan is the most expected by the observer as it accumulates the highest payoff for the job, given a light traffic condition. Notice that this plan evaluates to the highest $\mathcal{E}_\mathcal{A}$ score but is very inefficient given heavy traffic and hence in reality has the lowest payoff. Also, it conveys the wrong idea that there is light traffic as shown in Fig. 3 which conflicts with the ground truth. The active explicable plan (ActiveEXP) includes picking up the closest red guest G and three green guests (C, D and E). This plan has a slightly lower $\mathcal{E}_\mathcal{A}$ score but a much higher payoff than the EXP plan. More importantly, this plan conveys to the observer that there is heavy traffic as shown in Fig. 3. (c) by choosing to pick-up the green guests after picking up G. These results show that ActiveEXP is more efficient while still being explicable.



(a) (b) (c)

Fig. 3: (a) Dynamic belief update of all models in M for the Blocksworld. The bold line represents the belief update for the true robot model M_R. (b) Belief update of the robot's true model M_R at different noise levels of the observation model. (c) Traffic situation estimated by subjects with OP, EXP and ActiveEXP schedules where the ground truth is *heavy traffic*.

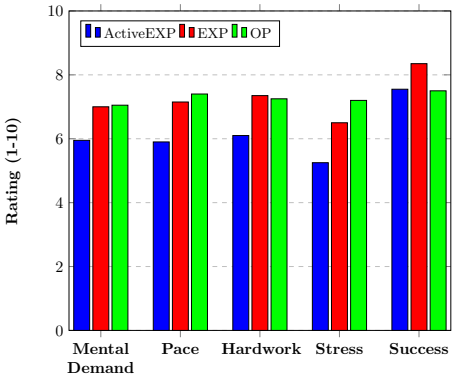


Fig. 5: NASA TLX study

Thus, they validate our two hypotheses $H1$ and $H2$.

In addition, this analysis is confirmed by the subjective results from our study presented in Fig. 5. using NASA TLX. We can see that ActiveEXP performed better than EXP and OP. From statistical analysis (independent t-test), we found that there is a significant difference between ActiveEXP and OP ($p = 0.015774$) and between ActiveEXP and EXP ($p = 0.063668$) at 0.10 level of significance. This result shows that ActiveEXP introduces less cognitive load than the EXP, which is somewhat surprising. This may be due to the fact that we are constantly expecting changes to occur. We will further analyze it in future work.

VI. CONCLUSION

In this paper, we introduced the problem of active explicable planning with dynamic modeling of the human's belief. It addressed a limitations of the existing work on explicable planning, which assumes a static human belief. We proposed a planning framework based on a Bayesian approach for generating active explicable plans. We evaluated our method against the existing planning methods and showed that an active explicable plan is more efficient without suffering explicability for human-robot teaming.

ACKNOWLEDGMENT

This research is supported in part by the NSF grants 1844524, 2047186, the NASA grant NNX17AD06G, and the AFOSR grant FA9550-18-1-0067.

REFERENCES

[1] Y. Zhang, S. Sreedharan, Anagha Kulkarni, T. Chakraborti, Hankui Zhuo, and S. Kambhampati. Plan explicability and predictability for robot task planning. *2017 ICRA*, pages 1313–1320, 2017.

[2] Anagha Kulkarni, T. Chakraborti, Yantian Zha, Satya Gautam Vadlamudi, Y. Zhang, and S. Kambhampati. Explicable robot planning as minimizing distance from expected behavior. *ArXiv*, abs/1611.05497, 2016.

[3] Anagha Kulkarni, Siddharth Srivastava, and S. Kambhampati. A unified framework for planning in adversarial and cooperative environments. In *AAAI*, 2019.

[4] T. Chakraborti, S. Kambhampati, Matthias Scheutz, and Y. Zhang. Ai challenges in human-robot cognitive teaming. *ArXiv*, abs/1707.04775, 2017.

[5] Mehrdad Zakershahrak, Akshay Sonawane, Ze Gong, and Yu Zhang. Interactive plan explicability in human-robot teaming. In *2018 RO-MAN*, pages 1012–1017. IEEE, 2018.

[6] T. Chakraborti, Anagha Kulkarni, S. Sreedharan, David E. Smith, and S. Kambhampati. Explicability? legibility? predictability? transparency? privacy? security? the emerging landscape of interpretable agent behavior. In *ICAPS*, 2019.

[7] S. Sreedharan, Anagha Kulkarni, T. Chakraborti, D. E. Smith, and S. Kambhampati. A bayesian account of measures of interpretability in human-ai interaction. *ArXiv*, abs/2011.10920, 2020.

[8] Anca D. Dragan and S. Srinivasa. Generating legible motion. In *Robotics: Science and Systems*, 2013.

[9] Aleck M. MacNally, N. Lipovetzky, M. Ramírez, and A. Pearce. Action selection for transparent planning. In *AAMAS*, 2018.

[10] M. Ramírez and H. Geffner. Probabilistic plan recognition using off-the-shelf classical planners. In *AAAI*, 2010.

[11] Henry A. Kautz and James F. Allen. Generalized plan recognition. In *AAAI*, 1986.

[12] Eugene Charniak and R. Goldman. A bayesian model of plan recognition. *Artif. Intell.*, 64:53–79, 1993.

[13] S. J. Levine and B. Williams. Concurrent plan recognition and execution for human-robot teams. In *ICAPS*, 2014.

[14] Ehab ElSalamouny, V. Sassone, and M. Nielsen. Hmm-based trust model. In *Formal Aspects in Security and Trust*, 2009.

[15] Y. Xia, Alexandre Parmentier, and R. Cohen. Trust modelling in dynamic environments. 2019.

[16] X. Liu and A. Datta. Modeling context aware dynamic trust using hidden markov model. In *AAAI*, 2012.

[17] Leslie Pack Kaelbling, Michael L Littman, and Anthony R Cassandra. Planning and acting in partially observable stochastic domains. *Artificial intelligence*, 101(1-2):99–134, 1998.

[18] Maria Fox and Derek Long. PDDL2.1: an extension to PDDL for expressing temporal planning domains. *CoRR*, abs/1106.4561, 2011.

[19] Chris Baker, Rebecca Saxe, and Joshua Tenenbaum. Bayesian theory of mind: Modeling joint belief-desire attribution. In *Proceedings of the annual meeting of the cognitive science society*, volume 33, 2011.